

NPS ARCHIVE  
1965.06  
MARTIN, J.

Doctor of Philosophy  
SOME BAYESIAN DECISION PROBLEMS  
IN A MARKOV CHAIN  
by  
JAMES JOHN MARTIN, Jr.  
Course I June 1965

Thesis  
M3576

Literary  
U.S. Naval Personnel  
Monterey, Calif.

35

**ACCOFILE SS**  
GENUINE PRICEBOARD BINDER  
CAT. NO. BF 2507 EMB

ACCO  
CHICAGO,  
LONDON,

ORDENSBURG, N.Y.  
TORONTO,  
MEXICO, D.F.

LIBRARY  
U. S. Naval Postgraduate School  
Monterey, California

SOME BAYESIAN DECISION PROBLEMS  
IN A MARKOV CHAIN

by

JAMES JOHN MARTIN, Jr.  
BA, University of Wisconsin  
(1955)  
MS, U. S. Naval Postgraduate School  
(1963)

Submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

at the

Massachusetts Institute of Technology

1965



Library  
U. S. Naval Postgraduate School  
Monterey, California

SOME BAYESIAN DECISION PROBLEMS  
IN A MARKOV CHAIN

by

JAMES JOHN MARTIN, Jr.

Submitted to the Department of Civil Engineering on May 14, 1965,  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy.

ABSTRACT

Some Bayesian decision models which involve a finite Markov chain with uncertain transition probabilities are studied in this report. The principal theoretical features of these models are set forth and various questions of numerical computation are considered.

It is assumed, for the most part, that the family of prior distributions of the matrix of transition probabilities is closed under sampling. This concept is defined and some properties of closed families of distributions are obtained. It is shown that there are an arbitrarily large number of such families, giving considerable generality to the entire study. A discounted adaptive control model for a Markov chain with alternative transition probabilities and rewards is then formulated as a set of functional equations. These equations are shown to have a unique bounded solution and a method of successive approximations is considered which converges monotonically to this solution.

The means, variances, and covariances of the n-step transition probabilities, the steady-state probabilities, the total discounted reward vector, and the process gain are then considered. It is shown that, under quite general conditions, the mean n-step transition probability matrix approaches the matrix of steady-state probabilities as  $n \rightarrow \infty$ . These results are applied to discounted terminal control models in which a Markov chain with alternative transition probabilities and rewards is sampled, at a cost, until a terminal decision point is reached. At that time a terminal policy is chosen and the system is operated indefinitely under this policy with no further sampling. It is shown that a terminal decision point is reached with probability one under an optimal sampling strategy. These models are formulated as functional equations, which are shown to have a unique bounded solution, and successive approximation techniques are investigated.

We then turn to fixed sample size analysis. The Whittle distribution, the matrix beta distribution, and the beta-Whittle distribution are introduced. It is assumed that a finite Markov chain with uncertain transition probabilities is observed for n consecutive transitions and the prior-posterior and preposterior analysis is developed.



The report concludes with a summary of specific results for a two-state Markov chain when the matrix of transition probabilities has the matrix beta distribution.

Thesis Supervisor: Ronald A. Howard

Title: Associate Professor of Electrical Engineering and  
Associate Professor of Industrial Management



#### ACKNOWLEDGEMENTS

Professor Ronald A. Howard, who first quickened my interest in Markov processes and Bayesian decision theory, supervised the research reported here. Early drafts of the report were read by Professor Gordon M. Kaufman, who made many helpful suggestions, and by Dr. George R. Murray and Ralph L. Miller. Professors E. Farnsworth Blisbee, John D. C. Little, and George P. Wadsworth also served as thesis readers.

The staff of the Massachusetts Institute of Technology Library System, in particular, the members of the Information and Reference Department, rendered valuable assistance. All computations were performed at the Computation Center of the Massachusetts Institute of Technology.

My entire graduate program was supported in full by the U. S. Navy, an organization to which I am deeply indebted.

Finally, I must acknowledge the valuable contribution of my wife, Betty, who carefully typed several versions of this study, the notation of which can at best be described as tortuous.



## TABLE OF CONTENTS

	Page
Title Page	1
Abstract	ii
Acknowledgments	iv
Table of Contents	v
Chapter 1. Introduction.	1
1.1 Bayesian Decision Theory and Markov Chains.	1
1.2 Definitions and Notation.	5
1.2.1 Markov Chain With Alternatives.	6
Chapter 2. Families of Distributions Closed Under Sampling.	9
2.1 Families of Distributions Closed Under a Sampling Rule.	9
2.2 The Matrix Beta Distribution.	12
2.3 Families of Distributions Closed Under the Consecutive Sampling Rule or the $\nu$ -Step Sampling Rule.	15
2.3.1 Families Closed Under Consecutive Sampling.	16
2.3.2 Families Closed Under $\nu$ -Step Sampling.	19
2.3.3 Large Sample Theory.	23
2.4 Some General Properties of Closed Families of Distributions.	32
Chapter 3. Adaptive Control Problems.	38
3.1 Discounted Processes. Formulation.	38
3.2 Existence and Uniqueness of $v_1(\Psi)$ .	46
3.3 Solution by Successive Approximations.	50



3.4 Recursive Computation.	53
3.5 Numerical Example.	60
3.6 Undiscounted Processes.	66
Chapter 4. Expected Steady-State Probabilities and Related Quantities.	63
4.1 The $n$ -Step Transition Probability Matrix.	69
4.2 The Steady-State Probability Vector.	73
4.2.1 Existence of the Moments of $\tilde{\pi}$ .	73
4.2.2 An Ergodic Theorem.	77
4.2.3 Successive Approximations.	86
4.2.4 Variance-Covariance Matrix of $\tilde{\pi}$ .	94
4.3 Expected Discounted Reward Vector.	96
4.3.1 Expected Value of $\tilde{Y}(\tilde{P})$ .	97
4.3.2 A Functional Equation for $\tilde{Y}(\tilde{P})$ .	98
4.3.3 A Numerical Example.	101
4.3.4 Variance-Covariance Matrix.	106
4.4 The Process Gain.	108
4.4.1 Mean and Variance of $\tilde{g}(\tilde{P})$ .	109
4.4.2 Approximations to the Mean and Covariance Matrix of $\tilde{Y}(\tilde{P})$ .	110
4.4.3 An Abelian Theorem.	113
Chapter 5. Terminal Control Problems.	116
5.1 Discounted Processes. Model I.	117
5.2 Existence and Uniqueness of Solutions. Successive Approximations.	123
5.3 Discounted Processes. Model II.	129
5.4 Approximate Terminal Decisions.	138
5.5 Undiscounted Processes.	144



5.5.1 Model III.	145
5.5.2 Model IV.	146
5.6 Discounted Processes With Set-Up Costs.	148
 Chapter 6. Distribution Theory.	150
6.1 The Whittle Distribution and Related Distributions.	150
6.1.1 The Whittle Distribution.	153
6.1.2 Moments of the Whittle Distribution.	154
6.1.3 The Whittle-1 Distribution.	161
6.1.4 The Whittle-2 Distribution.	162
6.1.5 Moments of the Whittle-2 Distribution.	165
6.2 The Multivariate Beta Distribution.	169
6.2.1 The Multivariate Beta Density Function.	170
6.2.2 The Multivariate Beta Distribution Function.	174
6.2.3 The Nonstandardized Multivariate Beta Distribution.	176
6.2.4 Marginal and Conditional Distributions.	177
6.2.5 Moment Formulas.	179
6.3 The Matrix Beta Distribution.	180
6.4 Extended Natural Conjugate Distributions.	188
6.5 The Beta-Whittle Distribution.	191
6.5.1 The Beta-Whittle Distribution.	192
6.5.2 The Beta-Whittle-2 Distribution.	195
 Chapter 7. Fixed Sample Size Analysis.	199
7.1 Initial State Known.	199



7.1.1	Prior-Posterior Analysis.	200
7.1.2	Sampling Distributions and Preposterior Analysis.	200
7.2	Initial State Unknown.	202
7.2.1	Prior-Posterior Analysis.	203
7.2.2	Sampling Distributions and Preposterior Analysis.	203
7.3	System Operating in the Steady State.	205
7.3.1	Prior-Posterior Analysis.	205
7.3.2	Sampling Distributions and Preposterior Analysis.	207
Chapter 8.	Specific Results for a Two-State Markov Chain.	209
8.1	Preliminaries.	209
8.2	Hypergeometric Coefficients.	211
8.3	Expected n-Step Transition Probabilities.	212
8.4	Expected Value of $p_{\alpha\beta}^{(\nu)}, p_{\gamma\delta}^{(\lambda)}$ .	216
8.5	Steady-State Probabilities.	217
8.6	Process Gain.	220
8.7	Total Discounted Reward Vector.	220
8.8	A Generalization of a Result of Sher.	223
Chapter 9.	Concluding Remarks.	229
Appendix A.	Glossary of Symbols.	233
Appendix B.	Program VITERATION To Solve Equations (3.2.1) and (3.2.2).	240
Appendix C.	Program PHI MATRIX To Compute Equation (4.1.2).	243



Appendix D. Program PIAPROX To Compute Equations (4.2.42).	245
Appendix E. Program VASIMP To Compute Equation (4.3.13).	248
Bibliography	251
Biographical Note	256



## CHAPTER 1

### INTRODUCTION

#### 1.1 Bayesian Decision Theory and Markov Chains.

The basic concept of a Markov chain was introduced by A. A. Markov in 1907 and since that time the literature on the subject has grown remarkably. Fundamental investigations by Kolmogorov in the 1930's extended the mathematical theory to chains with an infinite number of states; Doeblin and Doob made important contributions during the period 1935-1945. The present state of the theory of Markov chains is summarized by Chung [12].

By 1950 it was well recognized that the Markov chain is a useful model for a multitude of physical processes and an increasing number of applications of the mathematical theory have been made to problems in such fields as physics, chemistry, biology, and operations research. In these applications it is generally assumed that the matrix of transition probabilities is known, although, since 1954, questions of hypothesis-testing and maximum-likelihood estimation have been investigated. These latter results are summarized by Billingsley [10], who gives extensive references.

During the past two decades Savage's interpretation of the work of de Finetti on subjective probability has renewed interest in Bayesian decision theory. Contributions in this area have been made by many researchers, including Von Neumann, Wald, Blackwell, and Girshick, leading to the current work of Raiffa and Schlaifer [33], which, to a large degree,



presents a unified theory of statistical decisions which is suitable for applications.

Recent research at the Massachusetts Institute of Technology [13, 14, 38] has been directed toward the application of Bayesian decision theory to various models based on Markov chains with uncertain transition probabilities. These efforts have demonstrated both the feasibility of such decision models and the need for a more thorough investigation of the underlying mathematical theory. The present work attempts to establish a theoretical basis for some decision models which involve a finite Markov chain with uncertain transition probabilities; particular attention is given to sequential decision models. While we have dealt, for the most part, with matters of existence and convergence, the question of numerical computation has not been neglected. There are, however, many problems of numerical computation in this area which are yet to be solved.

In 1953, L. S. Shapley [36], using a game-theoretic formulation, studied one of the earliest sequential decision models in a Markov chain with alternative transition probabilities, which were assumed to be known. Similar game formulations have been examined more recently by Zachrisson [42] and Shor [37]. A more general class of Markovian decision models with known transition probabilities have been investigated by Blackwell [11], Derman [16], Howard [22], and others, using the techniques of linear and dynamic programming. These models have been extended to semi-Markov processes by Howard [23] and Jewell [24, 25]. Further references are given by Jewell [25].

Silver [38] has investigated various questions in a Markov chain with uncertain transition probabilities and rewards. In particular, he has treated the problem of a natural conjugate distribution for the



data-generating process of a Markov chain and has attempted to find the expected value of certain functions of the transition probabilities, such as the steady-state probability vector. These results assumed a specific prior distribution for the transition probabilities, a generalization of the beta distribution which we shall call the matrix beta distribution.\* Many of Silver's results are generalized in the present work.

Cozzolino [13] has examined a sequential decision model involving a two-state chain with uncertain transition probabilities. In a related study, Cozzolino, Gonzalez-Zubieta, and Miller [14] have developed heuristic methods for treating sequential decisions in a Markov chain with uncertain transition probabilities. Their findings are based on Monte Carlo studies.

The results of the present study are obtained under the assumption that the prior distribution function of the matrix of transition probabilities belongs to a family of distributions which is closed under consecutive sampling. This concept is formally defined in Chapter 2, where some properties of such families of distributions are derived. In particular, it is shown that there are an arbitrarily large number of such families, thus providing considerable generality to the entire study. Additional generality is obtained by stating all theorems in terms of distribution functions and Riemann-Stieltjes integrals, making them applicable to both discrete and continuous prior distributions..

In Chapter 3 we consider a discounted adaptive control model in which alternative transition probabilities in a Markov chain are sampled over an infinite time period. The problem of choosing a sequence of policies which maximizes the expected discounted reward over an infinite period

---

\* Cf. Section 6.3



is formulated in terms of a set of functional equations. It is shown that these equations have a unique solution and a method of successive approximations which converges monotonically to this solution is considered.

Certain functions of the transition probabilities, such as the n-step transition probabilities, the steady-state probabilities, the discounted total reward, and the gain, are treated in Chapter 4, where we obtain recursive equations for the means, variances, and covariances of these quantities. An important result of this chapter is a proof that, under quite general conditions, the mean n-step transition probability matrix approaches the matrix of mean steady-state probabilities as  $n \rightarrow \infty$ .

These results are applied in Chapter 5, where discounted and undiscounted terminal control models are studied. In these models of a Markov chain with alternative transition probabilities the decision-maker can sample various alternatives by paying a sampling cost. After a certain amount of information about the process is gained in this manner it becomes profitable for him to cease sampling and to choose a policy under which the system operates indefinitely. These models are formulated as functional equations and it is shown that, with probability one, a terminal decision point is reached under an optimal sampling strategy. We then show that there exists a unique solution to the functional equations and investigate a method of successive approximation.

The results of the first six chapters are obtained for any prior distribution function which belongs to a family closed under consecutive sampling. In Chapters 6-8 we consider a specific distribution for the transition probabilities which we call the matrix beta distribution. The distribution is defined in Chapter 6 and its main properties are derived. We also introduce, in this chapter, the Whittle distribution and the



beta-Whittle distribution. These probability distributions are utilized in Chapter 7, where we do prior-posterior and preposterior analysis for a Markov chain which is observed under the consecutive sampling rule. The transition count is identified as a sufficient statistic and is shown to have the Whittle distribution, conditional on a fixed value of the transition probability matrix. The natural conjugate distribution for this data-generating process is the matrix beta and the unconditional distribution of the transition count is the beta-Whittle distribution.

In Chapter 8 we consider the results of Chapters 2-6 in the case of a two-state Markov chain when the prior distribution of the transition probabilities is matrix beta. Explicit formulas for the expected values of various functions of the transition probabilities are given in terms of the parameters of the matrix beta distribution.

The results of this study are summarized in Chapter 9 and areas for future research are discussed.

## 1.2 Definitions and Notation.

The matrix with generic element  $p_{ij}$  is denoted  $\underline{P} = [p_{ij}]$ ; the row vector with generic element  $p_i$  is written  $\underline{p} = (p_1, \dots, p_N)$ . The matrix  $\underline{P}^t$  is the transpose of  $\underline{P}$ .

A vector  $\underline{x} = (x_1, \dots, x_N)$ , is a point in the  $N$ -dimensional Euclidean space,  $E_N$ , and we shall use the customary norm, or distance function,  $\|\underline{x}\|$ , defined by

$$\|\underline{x}\| = \left[ \sum_{i=1}^N x_i^2 \right]^{\frac{1}{2}}. \quad (1.7.1)$$

Similarly, the  $M \times N$  matrix  $\underline{P}$  is a point in  $E_{MN}$  and has the norm



$$\|\tilde{P}\| = \left[ \sum_{i=1}^M \sum_{j=i}^N p_{ij}^2 \right]^{\frac{1}{2}}. \quad (1.2.2)$$

Random quantities are denoted by the tilde; thus,  $\tilde{\Gamma}$ ,  $\tilde{p}$ ,  $\tilde{p}_{ij}$  are, respectively, a random matrix, a random vector, and a random variable.

Let  $h(\underline{P})$  be a scalar function of the  $M \times N$  matrix  $\underline{P}$ . Assume that each row of  $\underline{P}$  is subject to the constraint

$$\sum_{j=1}^N p_{ij} = 1. \quad i=1, \dots, M \quad (1.2.3)$$

If  $F(\underline{P})$  is a distribution function, the Riemann-Stieltjes integral of  $h$  is to be interpreted as an  $!!(l-1)$ -fold iterated over the independent elements of  $\underline{P}$ :

$$\int h(\underline{P}) dF(\underline{P}) = \quad (1.2.4)$$

$$\int \dots \int h(p_{11}, \dots, p_{1,N-1}, p_{21}, \dots, p_{M,N-1}) dF(p_{11}, \dots, p_{M,N-1}).$$

If  $\underline{h}(\underline{P}) = [h_{ij}(\underline{P})]$  is a matrix-valued function of  $\underline{P}$ , the Riemann-Stieltjes integral of  $\underline{h}$  is to be interpreted as the matrix of the integrals of each element:

$$\int \underline{h}(\underline{P}) dF(\underline{P}) = \left[ \int h_{ij}(\underline{P}) dF(\underline{P}) \right]. \quad (1.2.5)$$

**1.2.1 Markov Chain With Alternatives.** When we refer to a Markov chain with alternatives we mean the following process. Let there be  $N$  states which the system can occupy. Then the system is in state  $i$ , the decision-maker can choose one of  $K_i$  alternative transition vectors,  $\underline{p}_i^k = (p_{i1}^k, \dots, p_{iN}^k)$ , where  $p_{ij}^k$  is the probability that the system makes a transition to state  $j$ , given that it is currently in state  $i$  and the  $k$ th



alternative is used. The vectors  $\mu_i^k$  are stochastic vectors; that is,

$$p_{ij}^k \geq 0, \quad k=1, \dots, K_i \quad i, j = 1, \dots, N \quad (1.7.6a)$$

$$\sum_{j=1}^N p_{ij}^k = 1. \quad k=1, \dots, K \quad i=1, \dots, n \quad (1.2.6c)$$

With each transition vector,  $D_1^k$ , is associated a reward vector,  
 $x_1^k = (r_{11}^k, \dots, r_{1N}^k)$ , where  $r_{ij}^k$  is the reward earned when the system makes a transition from state  $i$  to state  $j$  under the  $k$ th alternative ( $-\infty < r_{ij}^k < \infty$ ,  $k=1, \dots, K_1$ ;  $i, j=1, \dots, N$ ).

The transition vectors can be arranged in a  $K \times H$  matrix,  $\mathbb{P}$ , where

$$K = \sum_{\{z\}}^N K_z :$$

$$\underline{\Phi} = \begin{bmatrix} 1 \\ p_1 \\ \vdots \\ \cdot^{K_1} \\ x_1 \\ \vdots \\ \cdot^{K_N} \\ p_N \end{bmatrix}.$$

Let the corresponding reward matrix be denoted by  $\underline{Q}$ . Reserving the term stochastic matrix for square matrices of non-negative elements whose rows sum to unity, we shall call a  $K \times N$  matrix whose elements satisfy (1.2.6) a generalized stochastic matrix.

A policy consists of the selection of one alternative in each state and may be expressed as a row vector  $\underline{\sigma} = (\sigma_1, \dots, \sigma_N)$ , where  $\sigma_i$  is the index of the alternative selected in state  $i$  ( $\sigma_i = 1, \dots, K_i$ ). The stochastic matrix which governs the transitions of the Markov chain under a specific policy,  $\underline{\sigma}$ , will be denoted by  $P(\underline{\sigma})$  or, if no confusion will result, by  $P$ . The corresponding reward matrix under policy  $\underline{\sigma}$  is  $R(\underline{\sigma})$ .



or  $\mathbb{R}$ . The set of all possible policy vectors,  $\Sigma$ , is denoted  $\Sigma$  and is a finite set.

The matrix  $\underline{\mathbb{P}}$  can be regarded as the parameter of a Markov chain with alternatives; uncertainty about  $\underline{\mathbb{P}}$  is expressed by regarding  $\underline{\mathbb{P}}$  as a random matrix with a prior distribution function,  $H(\underline{\mathbb{P}}|\psi)$ , which has the parameter  $\psi$ . In general,  $\psi$  is a point in a multidimensional Euclidean space. The range set of  $\underline{\mathbb{P}}$  is the set of all  $K \times N$  generalized stochastic matrices, denoted  $\mathcal{S}_{K,N}$ :

$$\mathcal{S}_{K,N} = \left\{ \underline{\mathbb{P}} \mid \underline{\mathbb{P}} \text{ is } K \times N, p_{ij}^k \geq 0, \sum_{j=1}^N p_{ij}^k = 1 \quad (k=1, \dots, K_1; i, j=1, \dots, N) \right\} \quad (1.2.8)$$

We remark that  $\mathcal{S}_{K,N}$  is a closed and bounded, hence, compact, subset of the  $KN$ -dimensional Euclidean space,  $E_{KN}$ . The distribution function,  $H(\underline{\mathbb{P}}|\psi)$ , is a function of the  $K(N-1)$  independent elements of  $\underline{\mathbb{P}}$ ,  $p_{11}^k, \dots, p_{1,N-1}^k$ , for  $k=1, \dots, K_1$  and  $i=1, \dots, N$ .  $H(\underline{\mathbb{P}}|\psi)$  has the usual properties of a multivariate distribution function; in particular,

$$\int_{\mathcal{S}_{K,N}} dH(\underline{\mathbb{P}}|\psi) = 1. \quad (1.2.9)$$

From  $H(\underline{\mathbb{P}}|\psi)$  can be obtained the marginal distributions of the  $\prod_{i=1}^{K_1} \mathbb{R}_+$  possible stochastic matrices,  $\tilde{\mathbb{P}}(\Sigma)$ . The marginal distribution function of  $\tilde{\mathbb{P}}(\Sigma)$  is denoted  $F_{\Sigma}(\underline{\mathbb{P}}|\psi)$  or, when the dependence on  $\Sigma$  is clear, simply  $F(\underline{\mathbb{P}}|\psi)$ . The range set of  $\tilde{\mathbb{P}}$  is  $\mathcal{S}_N$ , the set of all  $N \times N$  stochastic matrices:

$$\mathcal{S}_N = \left\{ \underline{\mathbb{P}} \mid \underline{\mathbb{P}} \text{ is } N \times N, p_{ij} \geq 0, \sum_{j=1}^N p_{ij} = 1 \quad (i, j=1, \dots, N) \right\}. \quad (1.2.10)$$



CHAPTER 2  
FAMILIES OF DISTRIBUTIONS CLOSED  
UNDER SAMPLING

Much of the discussion presented in the following chapters is carried out under the assumption that  $\Pi(\underline{\theta}|\underline{Y})$ , the prior distribution of  $\underline{\theta}$ , is a member of a family of distributions closed under a given sampling rule. We formally define this concept in the present chapter and derive some properties of such closed families of distributions which will be used in the sequel.

The notion of a family of distributions closed under sampling is, of course, not a new one. In Great Britain, G. A. Barnard [5] in 1954 and, more recently, G. B. Wetherill [40], have applied this concept to sampling inspection problems. In this country, R. Bellman [6] and Bellman and Kalaba [8] have used the idea in connection with adaptive control processes. A particular class of distributions closed under sampling, known as natural conjugate distributions, forms the basis of recent research by Raiffa and Schlaifer [33] in statistical decision theory.

The properties of closed families of distributions which are derived in this chapter and their application to decision problems in a Markov chain with alternatives are original with the present work.

#### 2.1 Families of Distributions Closed Under a Sampling Rule.

Consider a sequence of transitions within a Markov chain with alternatives. A sampling rule is a set of specifications which determine the following:



- a. The distribution of the initial state of the chain and the initial policy under which the process is operated.
- b. The transitions at which policy changes occur. These transitions may be determined probabilistically.
- c. The distribution of the new policy when a policy change occurs. This distribution is a probability mass function over the set of policies,  $\Sigma$ , and allows for randomized selection of policies.
- d. The transitions at which the state of the process is made known to the decision-maker. Those transitions may be determined probabilistically and, when they do occur, an observation of the process is said to have taken place. Thus, an observation of the process is a random variable whose range is the set of state indices,  $\{1, \dots, N\}$ .
- e. A rule for termination of sampling.

We adopt the convention that, if a policy change or an observation occurs at the  $n$ th transition, it takes place immediately after the  $n$ th transition has occurred.

There are two sampling rules which are of particular importance in succeeding chapters, consecutive sampling and  $\omega$ -stop sampling.

A consecutive sampling rule of size  $n$  is characterized as follows. A specific initial state and initial policy are selected with probability one. A total of  $n$  transitions are to occur, with  $n$  selected in advance. Each transition is observed. Policy changes, if they occur, take place at predetermined transitions and, at each change, a predetermined policy is chosen with probability one. Thus, a consecutive sampling rule of size  $n$  consists of  $n$  consecutive observations of the states of a Markov



chain with alternatives under a sequence of policies which is selected in advance of sampling.

A n-step sampling rule of size  $n$  may be described as follows. A positive integer,  $n$ , a sequence of  $n$  positive integers,  $\{v_1, \dots, v_n\}$ , and a sequence of  $n$  policies,  $\{\underline{\sigma}_1, \dots, \underline{\sigma}_n\}$ , are selected in advance of sampling. We allow the possibility that some or all of the  $\underline{\sigma}_i$  are equal. A specific initial state is chosen with probability one and a sequence of  $v_1$  transitions are allowed to occur under the policy  $\underline{\sigma}_1$ . The state of the Markov chain is observed after the  $v_1$ th transition. Then  $v_2$  transitions occur under policy  $\underline{\sigma}_2$ , the state being observed after the  $v_2$ th transition, and so on. A total of  $n$  observations are taken in this manner. The  $n$ -step sampling rule will be used in one of the terminal control models of Chapter 5.

We now proceed with the definition of a family of distributions closed under a sampling rule. A collection,  $\mathcal{H}$ , of probability distribution functions is said to be a family of distributions indexed by  $\Psi$  if all members of the collection have the same functional form and differ only in the values assigned to the parameter  $\Psi$ . The set of values which  $\Psi$  can assume is denoted  $\Psi$ , termed the admissible parameter set. The admissible parameter set is assumed to be a connected subset of a (possibly multidimensional) Euclidean space.

Let a sampling rule be specified and assume that a sample of  $n$  observations,  $\underline{x}_n = (x_1, \dots, x_n)$ , has resulted under that sampling rule. Denote by  $\ell(\underline{x}_n | \underline{\Psi})$  the likelihood of the sample  $\underline{x}_n$  under the given sampling rule given that  $\underline{\Psi} = \underline{\Psi}$ . Let the prior distribution function of  $\underline{\Psi}$  be  $\pi(\underline{\Psi} | \Psi)$ , a member of  $\mathcal{H}$ , a family of distributions indexed by  $\Psi$ . Then if  $d\pi(\underline{\Psi} | \Psi)$  is the prior probability that  $\underline{\Psi}$  lies in an infinitesimal



neighborhood of  $\underline{\Phi}$ , the posterior distribution function of  $\widetilde{\underline{\Phi}}$  is  $\Pi(\underline{\Phi} | \Psi, \underline{x}_n)$ , defined by means of Bayes' Theorem:

$$\text{d}\Pi(\underline{\Phi} | \Psi, \underline{x}_n) = \frac{\ell(\underline{x}_n | \underline{\Phi}) \text{d}\Pi(\underline{\Phi} | \Psi)}{\int \ell(\underline{x}_n | \underline{\Phi}) \text{d}\Pi(\underline{\Phi} | \Psi)} . \quad (2.1.1)$$

If  $\Pi(\underline{\Phi} | \Psi, \underline{x}_n) \in \Psi$  for all  $\Psi \in \Psi$  and all samples  $\underline{x}_n$  of non-zero probability, then  $\Psi$  is said to be closed with respect to the sampling rule which determines  $\ell(\underline{x}_n | \underline{\Phi})$ . In this case the posterior distribution is denoted  $\Pi(\underline{\Phi} | \Psi')$ , where

$$\Psi' = T(\Psi). \quad (2.1.2)$$

Here  $T$  is the mapping of  $\Psi$  into  $\Psi'$  induced by the transformation (2.1.1) when  $\Psi$  is closed under the given sampling rule.

In the special case where the sample consists of a single transition from state  $i$  to state  $j$  under the  $k$ th alternative in state  $i$ ,  $\Psi'$  will be written

$$\Psi' = T_{i,j}^k(\Psi). \quad (2.1.3)$$

If a fixed policy  $\underline{\pi}$  is in force, the superscript  $k = \sigma_j$  may be suppressed in (2.1.3).

In Section 2.3, families of distributions which are closed relative to the consecutive sampling and  $v$ -step sampling rules are discussed in detail. In order to carry out this discussion, some properties of the matrix beta distribution are required. Those properties are summarized in the next section.

## 2.2 The Matrix Beta Distribution.

The matrix beta density function, defined by equation (2.2.1) below, will be shown to be the natural conjugate distribution for the likelihood function of the consecutive sampling rule and, hence, is of intrinsic



importance. Moreover, as will be seen in Section 2.3, many of the properties of arbitrary families of distributions which are closed relative to the consecutive sampling rule or the v-step sampling rule are related to characteristics of the matrix beta distribution. For these reasons, the principal facts about this distribution are summarized in this section without proof. Complete derivations are given in Chapter 6.

The  $K \times N$  random generalized stochastic matrix,  $\underline{\underline{P}} = [\underline{p}_{ij}^k]$ , is said to have the matrix beta distribution with parameter  $\underline{\underline{m}} = [m_{ij}^k]$  if  $\underline{\underline{P}}$  has the joint density function

$$f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{m}}) = k(\underline{\underline{m}}) \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (\underline{p}_{ij}^k)^{m_{ij}^k} \Gamma(m_{ij}^k)^{-1}, \quad \underline{\underline{P}} \in \mathcal{S}_{K,N}$$

$$= 0. \quad \text{elsewhere} \quad (2.2.1)$$

The normalizing constant,  $k(\underline{\underline{m}})$ , is given by

$$k(\underline{\underline{m}}) = \prod_{i=1}^N \prod_{k=1}^{K_i} \frac{\Gamma(m_i^k)}{\prod_{j=1}^N \Gamma(m_{ij}^k)}, \quad (2.2.2)$$

where

$$m_i^k = \sum_{j=1}^N m_{ij}^k, \quad \begin{matrix} k=1, \dots, K_i \\ i=1, \dots, N \end{matrix} \quad (2.2.3)$$

The parameter  $\underline{\underline{m}}$  is a  $K \times N$  matrix such that

$$\underline{p}_{ij}^k > 0, \quad \begin{matrix} k=1, \dots, K_i \\ i,j=1, \dots, N \end{matrix} \quad (2.2.4)$$

It is shown in Chapter 6 that

$$\int_{\mathcal{S}_{K,N}} f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{m}}) d\underline{\underline{P}} = 1. \quad (2.2.5)$$

For  $k=1, \dots, K_i$  and  $i,j=1, \dots, N$ , the means and variances of the elements of  $\underline{\underline{P}}$  are given by the formulas

$$E[\underline{p}_{ij}^k] = \frac{m_{ij}^k}{m_i^k} = \bar{p}_{ij}^k \quad (2.2.6)$$



and

$$\begin{aligned} \text{var} [\bar{p}_{ij}^k] &= \frac{n_{ij}^k (M_1^k - n_{ij}^k)}{(n_{ij}^k)^2 (M_2^k + 1)} \\ &\approx \frac{\bar{p}_{ij}^k (1 - \bar{p}_{ij}^k)}{M_1^k + 1}. \end{aligned} \quad (2.2.7)$$

The covariances of the elements of  $\tilde{\underline{P}}$  are

$$\begin{aligned} \text{cov} [p_{\alpha\beta}^k, p_{\gamma\delta}^j] &= \frac{-n_{\alpha\beta}^k n_{\gamma\delta}^j}{(M_\alpha^k)^2 (M_\gamma^k + 1)}, & j, k = 1, \dots, K_1 \\ &= 0, & j \neq k \text{ or } \alpha \neq \gamma \end{aligned} \quad (2.2.8)$$

Let  $\underline{x}_n = (x_0, x_1, \dots, x_n)$  be a sample of  $n$  transitions observed under the consecutive sampling rule, where  $x_0$  is the initial state, known in advance of sampling. Let  $f_{ij}^k$  denote the number of transitions in  $\underline{x}_n$  from state  $i$  to state  $j$  under the  $k$ th alternative in state  $i$  ( $k=1, \dots, K_1$ ;  $i, j=1, \dots, N$ ) and define the transition count of the sample as the  $K \times N$  matrix  $\underline{F} = [f_{ij}^k]$ . Then the conditional probability, given that  $\tilde{\underline{P}} = \underline{P}$ , of observing the sample  $\underline{x}_n$  is

$$\prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_1} (p_{ij}^k)^{f_{ij}^k}. \quad (2.2.9)$$

If the rule by which the sample size  $n$  was selected is noninformative in the sense of Raiffa and Schlaiffer [33], then (2.2.9) is the likelihood of the sample  $\underline{x}_n$ . It is clear that  $\underline{F}$  is a sufficient statistic for this data-generating process and that the natural conjugate distribution is the matrix beta distribution.

Theorem 2.2.1 Let  $\tilde{\underline{P}}$  have the matrix beta distribution with parameter  $\underline{m}^*$  and suppose that a sample with transition count  $\underline{F}$  is observed under the consecutive sampling rule with noninformative stopping. Then



the posterior distribution of  $\tilde{\Phi}$  is matrix beta with parameter

$$\underline{m}'' = \underline{m}' + F. \quad (2.2.10)$$

**Proof.** By Bayes' Theorem the posterior distribution,  $D(\underline{\Phi} | \underline{m}', F)$ , is proportional to the product of the kernel of the likelihood function and the kernel of the prior distribution,

$$D(\underline{\Phi} | \underline{m}', F) \propto \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (p_{ij})^{n_{ij}^k} f_{ij}^{k-1} \quad (2.2.11)$$

The right side of (2.2.11) is the kernel of a matrix beta distribution with parameter  $\underline{m}' + F$ . Q.E.D.

**Corollary 2.2.2.** The family of matrix beta distributions is closed with respect to the consecutive sampling rule.

**Proof.** The corollary follows directly from Theorem 2.2.1. Q.E.D.

### 2.3 Families of Distributions Closed Under the Consecutive Sampling Rule or the v-Step Sampling Rule.

In the following chapters we shall confine our attention to models based on either the consecutive sampling rule or the v-step sampling rule. Some properties of families of distributions which are closed under either of these rules are established in this section. Specifically it is shown that there are an unlimited number of distinct families of distributions which are closed under the consecutive sampling rule, thus allowing the decision-maker considerable latitude in selecting a prior distribution for  $\tilde{\Phi}$ . A lemma of fundamental importance for the development of consecutive sampling models is next established. We then turn to families of distributions closed under the v-step sampling rule and it is shown that the class of such families is identical with the class of families consisting of probability mixtures of distributions from a family closed under consecutive sampling. It then follows that any family of



distributions closed under  $v$ -step sampling is also closed under consecutive sampling. Finally, it is proven that, for an arbitrary prior distribution on  $\underline{\Phi}$ , if  $n$  observations of the Markov chain are obtained under either sampling rule, then, with probability one, the probability mass of the posterior distribution tends to concentrate at  $\underline{Q}$ , the true state of nature, as  $n \rightarrow \infty$ .

### 2.3.1 Families Closed Under Consecutive Sampling.

In Section 2.2 it was shown that the natural conjugate distribution for the consecutive sampling rule is the matrix beta distribution. Extended natural conjugate distributions for this sampling rule can be constructed as follows. Let  $g(\underline{\Phi}|\omega)$  be a non-negative Borel function\* defined on  $\mathcal{S}_{K,N}$  which is positive over some subset of  $\mathcal{S}_{K,N}$ . The parameter  $\omega$  is a point belonging to  $\Omega$ , a subset of a Euclidean space. Let  $\underline{m} = [m_{ij}^k]$  be a  $K \times N$  matrix with

$$m_{ij}^k > 0, \quad \begin{matrix} k=1, \dots, K \\ i, j=1, \dots, N \end{matrix} \quad (2.3.1)$$

We assume that  $g(\underline{\Phi}|\omega)$  is sufficiently well-behaved that the integral

$$\int_{\mathcal{S}_{K,N}} \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (p_{ij}^k)^{m_{ij}^k - 1} g(\underline{\Phi}|\omega) d\underline{\Phi} = C(\underline{m}, \omega) \quad (2.3.2)$$

exists for all  $\omega \in \Omega$  and all  $\underline{m}$  which satisfy (2.3.1). Let

$$h(\underline{\Phi}|\underline{m}, \omega) = C(\underline{m}, \omega) \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (p_{ij}^k)^{m_{ij}^k - 1} g(\underline{\Phi}|\omega), \quad \underline{\Phi} \in \mathcal{S}_{K,N}$$

$$= 0. \quad \text{elsewhere} \quad (2.3.3)$$

The function  $h(\underline{\Phi}|\underline{m}, \omega)$  is a non-negative Borel function such that

$$\int_{\mathcal{S}_{K,N}} h(\underline{\Phi}|\underline{m}, \omega) d\underline{\Phi} = 1, \quad (2.3.4)$$

\* See Lévy [28], pp. 106, ff., for a discussion of Borel functions. A function which is continuous at all but a finite number of points can be shown to be a Borel function.



and is, therefore, a probability density function.

Corresponding to any function  $g(\underline{\Phi} | \omega)$  which satisfies the preceding requirements, we define the extended natural conjugate family,  $\mathcal{H}_g$ , indexed by the ordered pair  $(\underline{M}, \omega)$ , as the collection of probability density functions,  $h(\underline{\Phi} | \underline{M}, \omega)$ , defined by equation (2.3.3). The following theorem shows that  $\mathcal{H}_g$  is closed under the consecutive sampling rule.

Theorem 2.3.1 Let  $\mathcal{H}_g$  be a family of probability density functions,  $h(\underline{\Phi} | \underline{M}, \omega)$ , as defined by equation (2.3.3). If the prior distribution on  $\underline{\Phi}$  is  $h(\underline{\Phi} | \underline{M}^*, \omega^*) \in \mathcal{H}_g$  and if a sample  $\underline{x}_n = (x_0, \dots, x_n)$ , with transition count  $\underline{F} = [f_{ij}^k]$ , is observed by consecutive sampling, then the posterior distribution of  $\underline{\Phi}$  is  $h(\underline{\Phi} | \underline{M}^* + \underline{F}, \omega^*) \in \mathcal{H}_g$ . Thus,  $\mathcal{H}_g$  is closed under the consecutive sampling rule.

Proof. The posterior distribution of  $\underline{\Phi}$ ,  $D(\underline{\Phi} | \underline{M}^*, \omega^*, \underline{x}_n)$ , is proportional to the product of the kernel of the likelihood function and the kernel of the prior density function,

$$D(\underline{\Phi} | \underline{M}^*, \omega^*, \underline{x}_n) \propto \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (p_{ij}^k)^{n_{ij}} f_{ij}^{k-1} g(\underline{\Phi} | \omega^*), \quad (2.3.5)$$

from which the theorem follows. Q.E.D.

The parameter  $\omega$  provides the decision-maker with additional flexibility in encoding his prior knowledge about  $\underline{\Phi}$ . It is to be noted, however, that  $\omega$  remains unchanged in the posterior distribution and is, in that sense, a nuisance parameter. An example of an extended natural conjugate distribution is presented in Section 6.4.

The next result is of fundamental importance for the development of the succeeding chapters. Some additional notation is required. Let  $\underline{y} = (y_1, \dots, y_K)$  be a point in the Euclidean space  $E_K$  and let  $I$  denote an interval in  $E_K$ ,



$$I = \{x | a_i \leq x_i \leq b_i, (i = 1, \dots, KN)\}, \quad (2.3.6)$$

where  $a_i < b_i$  ( $i = 1, \dots, KN$ ). Let  $Q$  be a partition of  $I$  into a finite number of mutually exclusive and exhaustive intervals,  $I_1, \dots, I_n$ . For each  $I_v$ , we define the volume

$$v(I_v) = \prod_{i=1}^{KN} (b_i - a_i), \quad v = 1, \dots, n \quad (2.3.7)$$

and let  $v = \max_v \{v(I_v)\}$ . Finally, let  $x_{ij}^k$  denote the event that a transition occurs from state  $i$  to state  $j$  under the  $k$ th alternative in state  $i$ .

**Lemma 2.3.2** Let  $\mathcal{H}(\underline{\Phi}|\Psi) \in \mathcal{A}$ , a family of distributions closed under the consecutive sampling rule and let  $g(\underline{\Phi})$  be any integrable function of  $\underline{\Phi}$  defined on  $\mathcal{S}_{K,N}$ . Then the following identity is valid:

$$\int_{\mathcal{S}_{K,N}} p_{ij}^k g(\underline{\Phi}) dH(\underline{\Phi}|\Psi) = \bar{p}_{ij}^k(\Psi) \int_{\mathcal{S}_{K,N}} g(\underline{\Phi}) dH(\underline{\Phi}|T_{ij}^k(\Psi)), \quad (2.3.8)$$

$\Psi \in \mathbb{P}$   
 $k=1, \dots, K$   
 $i, j=1, \dots, i_N$

where  $\bar{p}_{ij}^k(\Psi)$  is the marginal expectation of  $p_{ij}^k$ .

**Proof.** Let  $I$  be an interval in  $I_{KN}$  which contains  $\mathcal{S}_{K,N}$ . For any partition  $Q$  of  $I$ , let  $\underline{\Phi}_v = [(p_{ij}^k)_v]$  denote an arbitrary point of  $I_v \cap \mathcal{S}_{K,N}$  and let  $\Delta_v(\Psi) = P[\widetilde{\Phi} \in I_v \cap \mathcal{S}_{K,N} | \Psi]$  when  $\widetilde{\Phi}$  has the distribution function  $H(\underline{\Phi}|\Psi)$ . Then

$$\int_{\mathcal{S}_{K,N}} p_{ij}^k g(\underline{\Phi}) dH(\underline{\Phi}|\Psi) = \lim_{\substack{n \rightarrow \infty \\ v \rightarrow 0}} \sum_{v=1}^n (\bar{p}_{ij}^k)_v g(\underline{\Phi}_v) \Delta_v(\Psi). \quad (2.3.9)$$

Using Bayes' Theorem, we have

$$\Delta_v(T_{ij}^k(\Psi)) = P[\widetilde{\Phi} \in I_v \cap \mathcal{S}_{K,N} | X_{ij}^k, \Psi]$$



$$= \frac{P[X_{ij}^k | \underline{\Phi} \in I_v \cap \mathcal{S}_{K,N}, \psi] \Delta_v(\psi)}{P[X_{ij}^k | \psi]} \quad (2.3.10)$$

But

$$P[X_{ij}^k | \underline{\Phi} \in I_v \cap \mathcal{S}_{K,N}, \psi] = \int_{I_v \cap \mathcal{S}_{K,N}} p_{ij}^k \frac{dH(\underline{\Phi} | \psi)}{\Delta_v(\psi)} \quad (2.3.11)$$

and, by the mean value theorem, there is a point  $\underline{\Phi}_v^* = [(p_{ij}^k)_v^*]$  of  $I_v \cap \mathcal{S}_{K,N}$  such that

$$P[X_{ij}^k | \underline{\Phi} \in I_v \cap \mathcal{S}_{K,N}, \psi] = (p_{ij}^k)_v^*. \quad (2.3.12)$$

Since  $\underline{\Phi}_v$  is an arbitrary point of  $I_v \cap \mathcal{S}_{K,N}$ , we may set  $(p_{ij}^k)_v = (p_{ij}^k)_v^*$  in (2.3.9). Then, noting that  $P[X_{ij}^k | \psi] = \bar{p}_{ij}^k(\psi)$ , equation (2.3.10) yields

$$(p_{ij}^k)_v^* \Delta_v(\psi) = \bar{p}_{ij}^k(\psi) \Delta_v(T_{ij}^k(\psi)) \quad (2.3.13)$$

and equation (2.3.9) becomes

$$\begin{aligned} \int_{\mathcal{S}_{K,N}} p_{ij}^k g(\underline{\Phi}) dH(\underline{\Phi} | \psi) &= \bar{p}_{ij}^k(\psi) \lim_{\substack{n \rightarrow \infty \\ v \rightarrow 0}} \sum_{v=1}^n g(\underline{\Phi}_v^*) \Delta_v(T_{ij}^k(\psi)) \\ &= \bar{p}_{ij}^k(\psi) \int_{\mathcal{S}_{K,N}} g(\underline{\Phi}) dH(\underline{\Phi} | T_{ij}^k(\psi)). \end{aligned} \quad (2.3.14)$$

Q.E.D.

### 2.3.2 Families Closed Under n-step Sampling.

Let us now consider the likelihood function associated with a  $v$ -step sampling rule of size  $n$ . This sampling rule is described by the sequence of transition numbers,  $\{v_1, \dots, v_n\}$ , and by the sequence of policies,  $\{\Sigma_1, \dots, \Sigma_n\}$ . Let  $\underline{x}_n = (x_0, \dots, x_n)$  denote the resulting observations, where  $x_0$  is the known initial state. Letting  $p_{ij}^{(v)}(\Sigma)$  denote the  $(i,j)$ th element of the matrix  $(P(\Sigma))^v$ , the conditional



probability, given that  $\underline{\Phi} = \underline{\Phi}$ , of observing the sample  $\underline{x}_n$ , is

$$p_{x_0 x_1}^{(v_1)} (\underline{\sigma}_1) p_{x_1 x_2}^{(v_2)} (\underline{\sigma}_2) \dots p_{x_{n-1} x_n}^{(v_n)} (\underline{\sigma}_n) = \prod_{j=1}^n p_{x_{j-1} x_j}^{(v_j)} (\underline{\sigma}_j). \quad (2.3.15)$$

If the rule by which the sample size  $n$  was chosen is noninformative, then (2.3.15) is the likelihood of the sample  $\underline{x}_n$ .

Let  $\mathcal{H}$  be a family of probability distributions indexed by  $\gamma \in \Psi$ . For any fixed positive integer  $m$  let  $a = (a_1, \dots, a_m)$  be a stochastic vector. A probability mixture of distributions from  $\mathcal{H}$  is defined to be the weighted sum

$$\Pi^*(\underline{\Phi} | \gamma_1, \dots, \gamma_m, a) = \sum_{i=1}^m a_i \Pi(\underline{\Phi} | \gamma_i), \quad (2.3.16)$$

where  $\Pi(\underline{\Phi} | \gamma_i) \in \mathcal{H}$  ( $i=1, \dots, m$ ). It is clear from the definition that  $\Pi^*(\underline{\Phi} | \gamma_1, \dots, \gamma_m, a)$  is also a probability distribution function for  $\widetilde{\Phi}$ . The mixed extension of  $\mathcal{H}$  is defined to be  $\mathcal{H}^*$ , the family of all probability mixtures of distributions from  $\mathcal{H}$  as  $a$  ranges over  $S_{1,m}$ , the set of  $m$ -dimensional stochastic vectors (for fixed  $m$ ), and as  $m$  ranges over the positive integers. Since  $\Pi(\underline{\Phi} | \gamma)$  is trivially a probability mixture,  $\mathcal{H} \subset \mathcal{H}^*$ .

The following theorems establish that a family of distributions is closed under  $v$ -step sampling if and only if it is the mixed extension of a family closed under consecutive sampling and that such a mixed extension is also closed under the consecutive sampling rule.

**Theorem 2.3.3** Let  $\mathcal{H}$  be a family of distributions closed under consecutive sampling and let  $\mathcal{H}^*$  be its mixed extension. Then  $\mathcal{H}^*$  is also closed under consecutive sampling.

**Proof.** Let  $\underline{x}_n$  denote a sample of size  $n$  obtained by consecutive sampling. If the prior distribution of  $\widetilde{\Phi}$  is  $\Pi^*(\underline{\Phi} | \gamma_1, \dots, \gamma_m, a) \in \mathcal{H}^*$ ,



where  $\underline{a}^* = (a_1^*, \dots, a_m^*)$ , and if  $\ell(\underline{x}_n | \underline{\varphi})$  is the likelihood function, then the posterior distribution is, using Bayes' Theorem and equation (2.3.16),

$$\begin{aligned} d\Pi^*(\underline{\varphi} | \psi_1^*, \dots, \psi_m^*, \underline{a}^*, \underline{x}_n) &= \frac{\ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_1^*, \dots, \psi_m^*, \underline{a}^*)}{\int_{\mathcal{S}_{K,N}} \ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_1^*, \dots, \psi_m^*, \underline{a}^*)} \\ &= \frac{\sum_{i=1}^N a_i^* \ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_i^*)}{\sum_{j=1}^N a_j^* \int_{\mathcal{S}_{K,N}} \ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_j^*)} \\ &= \sum_{i=1}^N a_i'' \delta_{\psi_i^*} \quad (2.3.17) \end{aligned}$$

where

$$a_i'' = \frac{a_i^* \int_{\mathcal{S}_{K,N}} \ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_i^*)}{\sum_{j=1}^N a_j^* \int_{\mathcal{S}_{K,N}} \ell(\underline{x}_n | \underline{\varphi}) d\Pi^*(\underline{\varphi} | \psi_j^*)} \quad i=1, \dots, N \quad (2.3.18)$$

and  $\psi_i''$  is defined by equation (2.1.2). Since  $\underline{a}'' = (a_1'', \dots, a_N'')$  is a stochastic vector, the posterior distribution of  $\widetilde{\underline{\varphi}}$  is

$$\pi^*(\underline{\varphi} | \psi_1'', \dots, \psi_N'') \in \mathbb{M}^*. \text{ Q.E.D.}$$

**Theorem 2.3.4** Let  $\mathbb{M}^*$  be a family of probability distributions indexed by  $\psi^* \in \Psi^*$ . A necessary and sufficient condition that  $\mathbb{M}^*$  be closed under the  $v$ -step sampling rule is that  $\mathbb{M}^*$  be the mixed extension of  $\mathbb{M}$ , a family of distributions closed under consecutive sampling.

**Proof.** First assume that  $n \geq 1$ . Let  $x_{ij}(v, \underline{\sigma})$  denote the observation of a transition from  $i$  to  $j$  over a transition interval of length  $v$  under the policy  $\underline{\sigma} = (\sigma_1, \dots, \sigma_N)$ . The likelihood function is



$$p_{ij}^{(v)}(\underline{\sigma}) = \sum_{i_1=1}^N \dots \sum_{i_{v-1}=1}^N p_{ii_1} p_{i_1 i_2} \dots p_{i_{v-1} j}, \quad (2.3.19)$$

which is the sum of  $N^{(v-1)}$  terms, each of which is the likelihood function for a sample sequence of length  $v$  observed under the consecutive sampling rule. Let  $\Pi^*(\underline{\theta} | \underline{\tau}^*) \in \mathcal{H}^*$  be the prior distribution of  $\underline{\theta}$ . The differential form of the posterior distribution has the kernel

$$d\Pi^*(\underline{\theta} | \underline{\tau}^*, x_{1j}(v, \underline{\sigma})) \propto \sum_{i_1=1}^N \dots \sum_{i_{v-1}=1}^N p_{ii_1} \dots p_{i_{v-1} j} d\Pi^*(\underline{\theta} | \underline{\tau}^*). \quad (2.3.20)$$

If  $\mathcal{H}^*$  is the mixed extension of a family closed under consecutive sampling, then Theorem 2.3.3 and equation (2.3.20) imply that  $\Pi^*(\underline{\theta} | \underline{\tau}^*, x_{1j}(v, \underline{\sigma})) \in \mathcal{H}^*$ . Moreover, if  $\mathcal{H}^*$  is not the mixed extension of a family of distributions closed under consecutive sampling, then  $p_{ii_1} \dots p_{i_{v-1} j} d\Pi^*(\underline{\theta} | \underline{\tau}^*)$  cannot be the kernel of a distribution in  $\mathcal{H}^*$  for all  $v$ ,  $\underline{\sigma}$ , and  $\underline{\tau}^*$ . Then, for some  $v$ ,  $\underline{\sigma}$ , and  $\underline{\tau}^*$ , the posterior distribution is a probability mixture of distributions, not all of which are in  $\mathcal{H}^*$ , and, therefore, the posterior distribution is not a member of  $\mathcal{H}^*$ . Thus, we have established necessity and sufficiency for the case  $n = 1$ .

For  $n > 1$ , the differential form of the posterior distribution of  $\underline{\theta}$  has the kernel

$$\begin{aligned} d\Pi^*(\underline{\theta} | \underline{\tau}^*, z_n) &\propto \prod_{j=1}^n p_{x_{j-1}, x_j}^{(v_j)} (\underline{\sigma}_j) d\Pi^*(\underline{\theta} | \underline{\tau}^*) \\ &\propto \prod_{j=2}^n p_{x_{j-1}, x_j}^{(v_j)} (\underline{\sigma}_j) d\Pi^*(\underline{\theta} | \underline{\tau}^*, x_{x_0 x_1}(v_1, \underline{\sigma}_1)), \end{aligned} \quad (2.3.21)$$

and the theorem follows by induction. Q.E.D.



Corollary 2.3.5 If  $\mathcal{H}^*$  is a family of distributions closed under the  $v$ -step sampling rule, then  $\mathcal{H}^*$  is also closed under the consecutive sampling rule.

Proof. The corollary follows immediately from Theorems 2.3.3 and 2.3.4. Q.E.D.

**2.3.3 Large Sample Theory.** Let  $\Pi(\underline{Q}|\mathcal{T})$  be an arbitrary prior distribution function of  $\widetilde{\mathcal{P}}$ . We now show that, if a sample of size  $n$  is observed under either the consecutive or the  $v$ -step sampling rule, the probability mass of the posterior distribution tends, as  $n \rightarrow \infty$ , to concentrate at  $\underline{Q}$ , the true state of nature, with probability one. This statement is made precise in Theorems 2.3.8 and 2.3.9. Not only are these results of interest on their own merits, but an important application of Theorems 2.3.8 and 2.3.9 will be made in Chapter 5, where the question of termination of sampling is considered for terminal control models.

Consider a sample of size  $n$  obtained under the  $v$ -step sampling rule. For a fixed state  $i$ , a fixed policy  $\underline{\Sigma}$ , and a fixed transition interval  $v$ , we shall say a trial occurs whenever the system makes a transition from state  $i$  to any other state over a transition interval of length  $v$  under the policy  $\underline{\Sigma}$ . For a fixed state  $j$ , let there be associated with the  $n$ th trial the random variable  $X_{n_i}(j)$  which takes the value 1 if the system is next observed in state  $j$  and the value zero otherwise. A sample of size  $n$  thus generates a sequence  $\{X_1(j), \dots, X_n(j)\}$  of independent, identically distributed random variables which, if  $\underline{Q}$  is the true state of nature, have the probability



function

$$P[\tilde{X}_a(j) = 1] = q_{ij}^{(v)}(\underline{\sigma}) \quad \begin{matrix} \text{as } i, 2, \dots \\ \text{j=1, ..., N} \end{matrix} \quad (2.3.22a)$$

$$P[\tilde{X}_a(j) = 0] = 1 - q_{ij}^{(v)}(\underline{\sigma}) \quad \begin{matrix} \text{as } i, 2, \dots \\ \text{j=1, ..., N} \end{matrix} \quad (2.3.22b)$$

and expected value

$$E[\tilde{X}_a(j)] = q_{ij}^{(v)}(\underline{\sigma}). \quad \begin{matrix} \text{as } i, 2, \dots \\ \text{j=1, ..., N} \end{matrix} \quad (2.3.23)$$

The following lemma is an immediate consequence of the strong law of large numbers.

**Lemma 2.3.6** Let  $(X_1(j), \dots, X_m(j))$  be an observation of size  $m$  of the sequence of trials defined above, for fixed states  $i$  and  $j$ , a fixed policy  $\underline{\sigma}$ , and a fixed transition interval  $v$ . If, as  $m \rightarrow \infty$ , state  $i$  is entered an infinite number of times and the policy  $\underline{\sigma}$  and transition interval  $v$  are used infinitely often when in state  $i$ , we have, with probability one,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{c=1}^m X_a(j) = q_{ij}^{(v)}(\underline{\sigma}), \quad j=1, \dots, N \quad (2.3.24)$$

where  $\underline{Q}$  is the true state of nature.

We remark that, if  $v=1$  and  $\sigma_i = k$ , Lemma 2.3.6 applies to the consecutive sampling rule and equation (2.3.24) becomes

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{c=1}^m X_a(j) = q_{ij}^k, \quad j=1, \dots, N \quad (2.3.25)$$

the limit holding with probability one.

A generalised stochastic matrix,  $\underline{Q} = [p_{ij}^k]$ , is said to be positive if all of its elements are positive, which implies that

$$0 < p_{ij}^k < 1. \quad \begin{matrix} k=1, \dots, K \\ i, j=1, \dots, N \end{matrix} \quad (2.3.26)$$



Lemma 2.3.7 Let  $x_n$  be a sample of size  $n$  obtained under the  $v$ -step sampling rule. Assume that, as  $n \rightarrow \infty$ , a fixed state  $i$  is observed infinitely often and that, when in state  $i$ , the policy  $\underline{\sigma}$  and transition interval  $v$  are used infinitely often. Then, if the true state of nature,  $\underline{Q}$ , is a positive matrix, every state  $j$  ( $j=1, \dots, N$ ) is, with probability one, observed infinitely often.

Proof. For fixed states  $i$  and  $j$ , the policy  $\underline{\sigma}$ , and the transition interval  $v$ , let  $\{x_m(j)\}$  be the sequence of trials generated by the sample  $x_n$ , as defined above. The hypotheses of the lemma imply that  $m \rightarrow \infty$  as  $n \rightarrow \infty$  and we have, by Lemma 2.3.6,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{a=1}^m x_a(j) = q_{ij}^{(v)}(\underline{\sigma}), \quad j=1, \dots, N \quad (2.3.27)$$

with probability one. Since  $q_{ij}^{(v)}(\underline{\sigma}) > 0$  for  $j=1, \dots, N$ , (2.3.27) implies that, with probability one,  $x_a(j) = 1$  infinitely often for each state  $j$ .

Q.E.D.

This lemma can probably be proved under the weaker assumption that  $\underline{Q}(\underline{\sigma})$  is ergodic, but it is sufficient for our purposes to assume that the true state of nature is a positive matrix. It will be shown in Chapter 4 that, for all prior distributions of  $\tilde{\Phi}$  which satisfy a mild continuity condition, the set of non-positive matrices is a set of measure zero.

We again remark that, by taking  $v = 1$ , Lemma 2.3.7 applies to samples obtained under the consecutive sampling rule as well as under the  $v$ -step sampling rule.

Let  $c$  be an arbitrary positive number and define  $\underline{\epsilon}$  to be the  $K \times N$  matrix each element of which is  $c$ . For any  $K \times N$  matrices  $\underline{\Phi}$  and  $\underline{Q}$ ,



say that

$$|\underline{P} - \underline{Q}| < \frac{\epsilon}{2} \quad (2.3.28)$$

if

$$\left| p_{ij}^k - q_{ij}^k \right| < \frac{\epsilon}{2}, \quad \begin{matrix} i=1, \dots, K_1 \\ i, j=1, \dots, N \end{matrix} \quad (2.3.29)$$

Clearly, if (2.3.28) holds, then

$$\begin{aligned} \|\underline{P} - \underline{Q}\| &= \left( \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^{K_i} (p_{ij}^k - q_{ij}^k)^2 \right)^{\frac{1}{2}} \\ &< \epsilon \sqrt{KN}, \end{aligned} \quad (2.3.30)$$

and the norm,  $\|\underline{P} - \underline{Q}\|$ , can be made arbitrarily small by an appropriate choice of  $\epsilon$ . Let  $\Pi(\underline{P} | \Psi)$  be an arbitrary prior distribution function of  $\underline{P}$  and assume that a sample,  $\underline{x}_n$ , of size  $n$  is observed. Denote by  $\Pi(\underline{P} | \Psi, \underline{x}_n)$  the posterior distribution of  $\underline{P}$  and, for fixed  $\underline{Q}$ , let

$$P_n [ |\tilde{\underline{P}} - \underline{Q}| < \frac{\epsilon}{2} ] = \int_{\underline{P} \in E} d\Pi(\underline{P} | \Psi, \underline{x}_n) \quad (2.3.31)$$

denote the posterior probability of the set

$$E = \{ \underline{P} \mid |\underline{P} - \underline{Q}| < \frac{\epsilon}{2} \} \subset \mathcal{A}_{K, n}. \quad (2.3.32)$$

When we say that the posterior probability mass tends, as  $n \rightarrow \infty$ , to concentrate at  $\underline{Q}$ , the true state of nature, with probability one, we mean that, for any  $\delta > 0$ ,

$$\lim_{n \rightarrow \infty} P_n [ |\tilde{\underline{P}} - \underline{Q}| < \frac{\epsilon}{2} ] = 1, \quad (2.3.33)$$

the limit holding with probability one.

Theorem 2.3.3 Let  $\Pi(\underline{P} | \Psi)$  be an arbitrary prior distribution function of  $\underline{P}$ . Let  $\underline{x}_n$  be a sample of size  $n$  obtained from a Markov chain with alternatives under the consecutive sampling rule. Assume that the sampling strategy is such that, as  $n \rightarrow \infty$ , if state  $i$  is entered infinitely often,



every alternative in state  $i$  is sampled infinitely often ( $i=1, \dots, N$ ).

If  $\underline{Q}$ , the true state of nature, is a positive matrix, then, for any  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P_n [ |\tilde{\underline{Q}} - \underline{Q}| < \epsilon ] = 1, \quad (2.3.34)$$

the limit holding with probability one, provided  $H(\underline{Q}|\psi)$  assigns positive probability to the set  $E$  defined by equation (2.3.32).

Proof. Let  $\underline{f}(n) = [f_{ij}^k(n)]$  be the transition count of the sample  $x_n$ . The posterior distribution of  $\tilde{\underline{Q}}$  is  $H(\underline{Q}|\psi, x_n)$ , where

$$dH(\underline{Q}|\psi, x_n) = \frac{\prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^K (p_{ij}^k)^{f_{ij}^k(n)} dH(\underline{Q}|\psi)}{\int_{\underline{Q}_{K,N}} \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^K (p_{ij}^k)^{f_{ij}^k(n)} dH(\underline{Q}|\psi)}. \quad (2.3.35)$$

Letting

$$\underline{m}(n) = [m_{ij}^k(n)] = [f_{ij}^k(n) + 1] \quad (2.3.36)$$

and multiplying the numerator and denominator of (2.3.35) by the normalizing constant  $k(\underline{m}(n))$  defined by equation (2.2.2), we have

$$dH(\underline{Q}|\psi, x_n) = \frac{f_{MB}^{(K,N)}(\underline{Q}|\underline{m}(n)) dH(\underline{Q}|\psi)}{\int f_{MB}^{(K,N)}(\underline{Q}|\underline{m}(n)) dH(\underline{Q}|\psi)}. \quad (2.3.37)$$

$$\text{Let } v_i^k(n) = \sum_{j=1}^N f_{ij}^k(n), \quad \begin{matrix} k=1, \dots, K \\ i=1, \dots, N \\ n=1, 2, \dots \end{matrix} \quad (2.3.38)$$

denote the number of times that alternative  $k$  is used in state  $i$  in a sample of size  $n$ . As  $n \rightarrow \infty$  at least one of the states of the chain is entered infinitely often. Lemma 2.3.7 and the hypotheses of the theorem imply that, with probability one, every state is entered infinitely often. Thus, under the assumed sampling strategy,  $v_i^k(n) \rightarrow \infty$  as  $n \rightarrow \infty$  with



probability one ( $k = 1, \dots, K_1; i = 1, \dots, N$ ). The mean of the distribution  $f_{MB}^{(K,N)}(\underline{\Phi}|\underline{M}(n))$  is  $\bar{\underline{\Phi}}(n) = [\bar{p}_{ij}^k(n)]$ , where

$$\bar{p}_{ij}^k(n) = \frac{\bar{p}_{ij}^k(n) + 1}{v_{ij}^k(n) + n}. \quad \begin{matrix} k=1, \dots, K_1 \\ i, j=1, \dots, N \\ n=1, 2, \dots \end{matrix} \quad (2.3.39)$$

Thus, if  $\underline{Q} = [q_{ij}^k]$ , Lemma 2.3.6 implies that, with probability one,

$$\lim_{n \rightarrow \infty} \bar{p}_{ij}^k(n) = q_{ij}^k > 0. \quad \begin{matrix} k=1, \dots, K_1 \\ i, j=1, \dots, N \end{matrix} \quad (2.3.40)$$

We now show that, as  $n \rightarrow \infty$ , the probability mass of  $f_{MB}^{(K,N)}(\underline{\Phi}|\underline{M}(n))$  tends, with probability one, to concentrate at  $\underline{Q}$ . If  $\widetilde{\underline{\Phi}}$  is a random matrix with the density function  $f_{MB}^{(K,N)}(\underline{\Phi}|\underline{M}(n))$ , the marginal variance of  $\bar{p}_{ij}^k$  is

$$v_{ij}^k(n) = \frac{\bar{p}_{ij}^k(n) [1 - \bar{p}_{ij}^k(n)]}{v_{ij}^k(n) + n + 1} \leq \frac{\frac{1}{4}}{v_{ij}^k(n) + n + 1}. \quad \begin{matrix} k=1, \dots, K_1 \\ i, j=1, \dots, N \\ n=1, 2, \dots \end{matrix} \quad (2.3.41)$$

Thus, with probability one,

$$\lim_{n \rightarrow \infty} v_{ij}^k(n) = 0. \quad \begin{matrix} k=1, \dots, K_1 \\ i, j=1, \dots, N \end{matrix} \quad (2.3.42)$$

Let  $\epsilon^* > 0$  and  $\delta$  ( $0 < \delta < 1$ ) be given. Define the set  $E \subset \mathcal{S}_{K,N}$  as in equation (2.3.32) and let

$$P[\|\widetilde{\underline{\Phi}} - \underline{Q}\| < \epsilon^* | \underline{M}(n)] = \int_{\underline{\Phi} \in E} f_{MB}^{(K,N)}(\underline{\Phi} | \underline{M}(n)) d\underline{\Phi}. \quad (2.3.43)$$

Since



Since

$$\left\{ \underline{\underline{P}} \mid |\underline{\underline{P}} - \underline{\underline{Q}}| < \frac{\epsilon^0}{n} \right\} = \bigcap_{i,j,k} \left\{ \underline{\underline{P}} \mid |p_{ij}^k - q_{ij}^k| < \epsilon^0 \right\},$$

De Morgan's law yields

$$\left\{ \underline{\underline{P}} \mid |\underline{\underline{P}} - \underline{\underline{Q}}| < \frac{\epsilon^0}{n} \right\}^C = \bigcup_{i,j,k} \left\{ \underline{\underline{P}} \mid |p_{ij}^k - q_{ij}^k| \geq \epsilon^0 \right\},$$

where C denotes the set complement.

$$1 - P [ |\underline{\underline{P}} - \underline{\underline{Q}}| < \frac{\epsilon^0}{n} | \mathcal{M}(n) ] \leq \sum_{i,j,k} P [ |p_{ij}^k - q_{ij}^k| \geq \epsilon^0 | \mathcal{M}(n) ] \quad (2.3.44)$$

Let

$$\mathcal{E}_{ij}^k = \left\{ \underline{\underline{P}} \mid |p_{ij}^k - q_{ij}^k| \geq \epsilon^0 \right\} \subset \mathcal{S}_{K,N} \quad \begin{matrix} k=1, \dots, K \\ i,j=1, \dots, n \end{matrix} \quad (2.3.45)$$

The marginal variance of  $\tilde{p}_{ij}^k$  is

$$\begin{aligned} v_{ij}^k(n) &= \int_{\mathcal{S}_{K,N}} (p_{ij}^k - \tilde{p}_{ij}^k(n))^2 f_{MP}^{(K,N)}(\underline{\underline{P}} | \mathcal{M}(n)) d\underline{\underline{P}} \\ &= -(\tilde{p}_{ij}^k(n) - q_{ij}^k)^2 + \int_{\mathcal{S}_{K,N}} (p_{ij}^k - q_{ij}^k)^2 f_{MP}^{(K,N)}(\underline{\underline{P}} | \mathcal{M}(n)) d\underline{\underline{P}} \end{aligned} \quad (2.3.46)$$

But

$$\begin{aligned} &\int_{\mathcal{S}_{K,N}} (p_{ij}^k - q_{ij}^k)^2 f_{MP}^{(K,N)}(\underline{\underline{P}} | \mathcal{M}(n)) d\underline{\underline{P}} \\ &\geq \int_{\underline{\underline{P}} \in \mathcal{E}_{ij}^k} (p_{ij}^k - q_{ij}^k)^2 f_{MP}^{(K,N)}(\underline{\underline{P}} | \mathcal{M}(n)) d\underline{\underline{P}} \\ &\geq \epsilon^0 \cdot P [ |\tilde{p}_{ij}^k - q_{ij}^k| \geq \epsilon^0 | \mathcal{M}(n) ] \end{aligned} \quad (2.3.47)$$

and, using (2.3.46),



$$P \left[ \left| \tilde{p}_{ij}^k - q_{ij}^k \right| \geq \epsilon^* \mid \underline{m}(n) \right] \leq \frac{1}{\epsilon^{*2}} \left[ v_{ij}^k(n) + (\tilde{p}_{ij}^k(n) - q_{ij}^k)^2 \right]. \quad (2.3.48)$$

$k=1, \dots, K$   
 $i, j=1, \dots, N$   
 $n=1, 2, \dots$

By equations (2.3.40) and (2.3.42), there exists an integer  $n^*$  such that, for all  $n > n^*$ ,

$$\frac{1}{\epsilon^{*2}} \left[ v_{ij}^k(n) + (\tilde{p}_{ij}^k(n) - q_{ij}^k)^2 \right] < \frac{6}{KN}, \quad k=1, \dots, K, \quad i, j=1, \dots, N \quad (2.3.49)$$

with probability one. Thus, for  $n > n^*$ , the inequality (2.3.44) becomes

$$1 - P \left[ |\underline{\Phi} - \underline{Q}| < \frac{\epsilon^*}{2} \mid \underline{m}(n) \right] < \delta, \quad n > n^* \quad (2.3.50)$$

and

$$P \left[ |\underline{\Phi} - \underline{Q}| < \frac{\epsilon^*}{2} \mid \underline{m}(n) \right] > 1 - \delta, \quad n > n^* \quad (2.3.51)$$

with probability one. Since  $\delta$  is arbitrary,

$$\lim_{n \rightarrow \infty} P \left[ |\underline{\Phi} - \underline{Q}| < \frac{\epsilon^*}{2} \mid \underline{m}(n) \right] = 1, \quad (2.3.52)$$

the limit holding with probability one.

Again defining  $E$  as in (2.3.32) and letting  $E^c$  be the complement of  $E$  in  $\mathcal{S}_{K,N}$ , we have, from equation (2.3.37),

$$P_n \left[ |\underline{\Phi} - \underline{Q}| < \frac{\epsilon^*}{2} \right] = \int_{\underline{\Phi} \in E} dH(\underline{\Phi} \mid \Psi, \underline{x}_n) \\ \int_{\underline{\Phi} \in E} f_{MB}^{(K,N)}(\underline{\Phi} \mid \underline{m}(n)) dH(\underline{\Phi} \mid \Psi) \quad (2.3.53)$$

---


$$\int_{\underline{\Phi} \in E} f_{MB}^{(K,N)}(\underline{\Phi} \mid \underline{m}(n)) dH(\underline{\Phi} \mid \Psi) + \int_{\underline{\Phi} \in E^c} f_{MB}^{(K,N)}(\underline{\Phi} \mid \underline{m}(n)) dH(\underline{\Phi} \mid \Psi)$$

Let  $\delta(n)$  be the maximum of the continuous function  $f_{MB}^{(K,N)}(\underline{\Phi} \mid \underline{m}(n))$  on the compact set  $E^c$ . Equation (2.3.52) implies that  $\lim_{n \rightarrow \infty} \delta(n) = 0$ , with probability one. Thus,



$$P_n [ |\underline{\hat{P}} - \underline{Q} | < \underline{\epsilon} ] \geq \frac{\int_{\underline{\epsilon} \in E} f_{MP}^{(K, N)} (\underline{\epsilon} | \underline{M}(n)) dH(\underline{\epsilon} | \psi)}{\delta(n) + \int_{\underline{\epsilon} \in E} f_{MP}^{(K, N)} (\underline{\epsilon} | \underline{M}(n)) dH(\underline{\epsilon} | \psi)} > 0 \quad (2.3.54)$$

and, with probability one,

$$\lim_{n \rightarrow \infty} P_n [ |\underline{\hat{P}} - \underline{Q} | < \underline{\epsilon} ] = 1.$$

Q.E.D.

**Theorem 2.3.9** Let  $H(\underline{\psi} | \psi)$  be an arbitrary prior distribution function of  $\underline{\hat{P}}$ . Let  $x_n$  be a sample of size  $n$  obtained from a Markov chain with alternatives under the  $v$ -step sampling rule. Assume that, when the system is observed in state  $i$ , the sampling rule is restricted to policies from  $\Sigma_i \subset \Sigma$  and to transition intervals from the finite set  $I_i = \{v_1, \dots, v_N\}$ , such that, as  $n \rightarrow \infty$ , if state  $i$  is observed infinitely often, every policy in  $\Sigma_i$  and every transition interval in  $I_i$  are used infinitely often ( $i = 1, \dots, K$ ). If  $\underline{Q}$ , the true state of nature, is a positive matrix, then for any  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P_n [ |\underline{\hat{P}} - \underline{Q} | < \epsilon ] = 1, \quad (2.3.55)$$

the limit holding with probability one, provided  $H(\underline{\psi} | \psi)$  assigns positive probability to the set  $E$  defined by equation (2.3.32).

**Proof.** Let  $K_i$  be the total number of ordered pairs,  $(\underline{\sigma}, v)$ , where  $\underline{\sigma} \in \Sigma_i$  and  $v \in I_i$  ( $i = 1, \dots, K$ ), and let  $K = \sum_{i=1}^K K_i$ . When in state  $i$ , let  $k$  index the possible policy and transition interval combinations,  $(\underline{\sigma}, v)$ . For  $\underline{\sigma} \in \Sigma_i$  and  $v \in I_i$ , let

$$\pi_{ij}^k = p_{ij}^{(v)}(\underline{\sigma}), \quad k=1, \dots, K \quad (2.3.56)$$

and define the  $K \times N$  matrix  $\underline{\Pi} = [\pi_{ij}^k]$ . Clearly,  $\underline{\Pi}$  is a generalized stochastic matrix. If the index  $k$  corresponds to the pair  $(\underline{\sigma}, v)$ , let



$f_{ij}^k(n)$  be the number of times a transition occurred from state  $i$  to state  $j$  in the sample  $\underline{x}_n$  over the transition interval  $v$  when the system was governed by the policy  $\underline{\Phi}$ . Then the posterior distribution of  $\widetilde{\Phi}$  is  $H(\underline{\Phi} | \underline{\gamma}, \underline{x}_n)$ , where

$$\begin{aligned} dH(\underline{\Phi} | \underline{\gamma}, \underline{x}_n) &= \frac{\prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (\pi_{ij}^k)^{f_{ij}^k(n)} dH(\underline{\Phi} | \underline{\gamma})}{\int_{\mathcal{S}_{K,N}} \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} (\pi_{ij}^k)^{f_{ij}^k(n)} dH(\underline{\Phi} | \underline{\gamma})} \\ &= \frac{f_{MP}^{(K,N)}(\underline{\Phi} | \underline{M}(n)) dH(\underline{\Phi} | \underline{\gamma})}{\int_{\mathcal{S}_{K,N}} f_{MP}^{(K,N)}(\underline{\Phi} | \underline{M}(n)) dH(\underline{\Phi} | \underline{\gamma})}, \end{aligned} \quad (2.3.57)$$

where  $\underline{M}(n) = [f_{ij}^k(n) + 1]$ . The proof of the theorem from this point is identical to the proof of Theorem 2.3.8. Q.E.D.

We remark that the assumptions in Theorems 2.3.8 and 2.3.9 concerning policies which are used infinitely often are not restrictive. It will usually be possible, after a finite amount of sampling, to eliminate from further consideration those policies which are used only a finite number of times. Examples of such elimination of policies by dominance arguments will be given in Chapter 5. In any case, the theorems apply to the marginal distribution of those alternative rows of  $\widetilde{\Phi}$  which are observed infinitely often.

## 2.4 Some General Properties of Closed Families of Distribution.

Let  $\mathcal{H}$  be a family of distributions indexed by  $\gamma \in \Gamma$  which is closed under an arbitrary sampling rule. Some general properties of  $\mathcal{H}$  are derived in this section. The symbol  $\ell(x_n | \underline{\Phi})$  will be used



throughout for the likelihood of a sample of size  $n$ , conditional on  $\tilde{\underline{\Phi}} = \underline{\Phi}$ , under the given sampling rule.

**Theorem 2.4.1** Let  $\tilde{\underline{\Phi}}$  have a discrete prior distribution,

$$P[\tilde{\underline{\Phi}} = \underline{\Phi}_1] = a_1, \quad \underline{\Phi}_1 \in \mathcal{S}_{K,N} \quad (2.4.1)$$

$$\text{is } i=1, 2, \dots, m$$

where  $a_i \geq 0$ ,  $\sum_{i=1}^m a_i = 1$ . For a fixed integer,  $n$ , let  $\mathcal{A}_n$  be the family of all such discrete distributions, indexed by  $\underline{g} = (a_1, a_2, \dots, a_n)$ . Then  $\mathcal{A}_n$  is closed under all sampling rules.

**Proof.** Let  $\ell(x_n | \underline{\Phi})$  be the likelihood function for an arbitrary sampling rule. If  $\underline{g}'$  is the prior distribution of  $\tilde{\underline{\Phi}}$ , the posterior probability of  $\underline{\Phi}_i$  is

$$a_i'' = \frac{\ell(x_n | \underline{\Phi}_i) a_i'}{\sum_{i=1}^m \ell(x_n | \underline{\Phi}_i) a_i'} \quad (2.4.2)$$

Since  $a_i'' \geq 0$  and  $\sum_{i=1}^m a_i'' = 1$ ,  $\underline{g}'' = (a_1'', a_2'', \dots, a_n'') \in \mathcal{A}_n$ .

Q.E.D.

This theorem, while almost trivial, is of considerable importance for the solution of Bayesian decision problems in a Markov chain in practice. In many cases it may be feasible to place positive probability on only a finite set of points of  $\mathcal{S}_{K,N}$  and to solve the corresponding discrete problem, thus considerably simplifying the computations.<sup>c</sup> We shall not emphasize this consideration any further since most of our theorems are stated in terms of Stieltjes integrals and, hence, are

---

<sup>c</sup> Cf. Silver [38], ch. 2.



applicable to discrete, continuous, and mixed prior distributions.

Theorem 2.4.2 Let  $\mathcal{H} = \{h(\underline{\theta}|\psi) | \psi \in \Psi\}$  be a family of distributions closed under a given sampling rule and, for a fixed policy  $\underline{\sigma}$ , let  $\mathcal{F}_{\underline{\sigma}} = \{F_{\underline{\sigma}}(P|\psi) | \psi \in \Psi^*\}$ , where  $\Psi^* \subset \Psi$ , be the corresponding family of marginal distributions of the  $N \times N$  stochastic matrix  $\tilde{P}(\underline{\sigma})$ . If the sampling rule is such that, for  $n=1, 2, \dots$ , it is possible to observe a sample of size  $n$  under the fixed policy  $\underline{\sigma}$ , and if the likelihood of any sample observed under the policy  $\underline{\sigma}$  does not depend on elements of  $\tilde{P}$  not in  $\tilde{P}(\underline{\sigma})$ , then  $\mathcal{F}_{\underline{\sigma}}$  is also closed under the given sampling rule.

Proof. Let  $\ell(x_n|P)$  be the likelihood function corresponding to the given sampling rule and let  $\ell_{\underline{\sigma}}(x_n|P)$  be the likelihood of the sample  $x_n$  from the Markov chain governed by  $\tilde{P}(\underline{\sigma})$ . The hypotheses of the theorem imply that

$$\ell_{\underline{\sigma}}(x_n|P) = \ell(x_n|P) \quad (2.4.3)$$

for all samples  $x_n$  from the chain governed by  $\tilde{P}(\underline{\sigma})$ . Let  $X_{\underline{\sigma}}$  be the range set of the  $(K-N) \times N$  generalized stochastic matrix formed by deleting from  $\tilde{P}$  all rows  $p_1^k$  such that  $k = \sigma_i$  ( $i=1, \dots, N$ ). Then, if  $F_{\underline{\sigma}}(P|\psi^*)$  is the marginal prior distribution of  $\tilde{P}(\underline{\sigma})$  and if the sample  $x_n$  is observed, the posterior distribution of  $\tilde{P}(\underline{\sigma})$  is  $F_{\underline{\sigma}}(P|\psi^*, x_n)$ , where

$$\begin{aligned} dF_{\underline{\sigma}}(P|\psi^*, x_n) &= \frac{\ell_{\underline{\sigma}}(x_n|P)dF_{\underline{\sigma}}(P|\psi^*)}{\int_N \ell_{\underline{\sigma}}(x_n|P)dF_{\underline{\sigma}}(P|\psi^*)} \\ &= \frac{\int_P \ell(x_n|P)dH(P|\psi^*)}{\int_{K,N} \ell(x_n|P)dH(P|\psi^*)} \end{aligned}$$



$$= \int_{\underline{K}_n} dH(\underline{\varPhi}|\psi^n) = dF_{\underline{\varPhi}}(\underline{P}|\psi^n) \quad (2.4.4)$$

for all  $\psi^n \in \Psi$ , where  $\psi^n$  is defined by equation (2.1.2). Thus,

$F_{\underline{\varPhi}}(\underline{P}|\psi^n, \underline{x}_n) \in \bar{\mathcal{F}}_n$  and  $\bar{\mathcal{F}}_n$  is closed under the sampling rule. Q.E.D.

The next theorem deals with the continuity of the expectation

$$\bar{g}(\psi) = \int g(\underline{\varPhi}) dH(\underline{\varPhi}|\psi)$$

when regarded as a function of  $\psi$ , where  $g(\underline{\varPhi})$  is any integrable function of  $\underline{\varPhi}$ .

A distribution function,  $H(\underline{\varPhi}|\psi)$ , is said to be continuous in  $\psi$  at a point  $\underline{\varPhi} \in \mathcal{S}_{K,N}$  if, for any  $\epsilon > 0$ , there exists a  $\delta > 0$  such that, for any fixed  $\psi$ ,  $|H(\underline{\varPhi}|\psi) - H(\underline{\varPhi}|\psi^*)| < \epsilon$  whenever  $\|\psi - \psi^*\| < \delta$ . A point,  $\underline{\varPhi}_0 \in \mathcal{S}_{K,N}$ , is said to be a continuity point of  $H(\underline{\varPhi}|\psi)$  if  $H(\underline{\varPhi}|\psi)$  is a continuous function of  $\underline{\varPhi}$  at  $\underline{\varPhi}_0$  for any fixed value of  $\psi$ .

Definition. Let  $\mathcal{H}$  be a family of distribution functions indexed by  $\psi \in \Psi$ .  $\mathcal{H}$  is said to be a family of distribution functions continuous in  $\psi$  if, whenever  $H(\underline{\varPhi}|\psi) \in \mathcal{H}$ ,  $H(\underline{\varPhi}|\psi)$  is a continuous function of  $\psi$  at each of its continuity points,  $\underline{\varPhi}$ .

Theorem 2.4.3 Let  $\mathcal{H}$  be a family of distribution functions indexed by  $\psi \in \Psi$  which is continuous in  $\psi$  and let  $g(\underline{\varPhi})$  be any integrable function of  $\underline{\varPhi}$  defined on a set  $S \subset \mathcal{S}_{K,N}$ . If  $\bar{g}(\psi)$  is a function of  $\psi$  defined by integral

$$\bar{g}(\psi) = \int_S g(\underline{\varPhi}) dH(\underline{\varPhi}|\psi), \quad \psi \in \Psi \quad (2.4.5)$$

then  $\bar{g}(\psi)$  is continuous on  $\Psi$ .



Proof. Let  $\gamma$  be fixed and let  $\{\gamma_n\}$  be any sequence of points of  $\Gamma$  which converges to  $\gamma$ , where  $\gamma_n \neq \gamma$  ( $n=1, 2, \dots$ ). Let  $H(\underline{P}|\gamma_n)$  be the corresponding sequence of distribution functions from  $\mathcal{H}$ . Since  $\mathcal{H}$  is continuous in  $\gamma$ ,

$$\lim_{n \rightarrow \infty} H(\underline{P}|\gamma_n) = H(\underline{P}|\gamma) \quad (2.4.6)$$

at every continuity point of  $H(\underline{P}|\gamma)$  and, by the Helly-Bray Theorem,<sup>\*</sup>

$$\lim_{n \rightarrow \infty} \int_S g(\underline{Q}) dH(\underline{P}|\gamma_n) = \int_S g(\underline{Q}) dH(\underline{P}|\gamma). \quad (2.4.7)$$

Thus, for every sequence  $\{\gamma_n\}$  which converges to  $\gamma$ ,  $\{\bar{g}(\gamma_n)\} \rightarrow \bar{g}(\gamma)$  and, therefore,  $\bar{g}(\gamma)$  is continuous at  $\gamma$ . Q.E.D.

Corollary 2.4.4 Let  $\mathcal{H} = \{H(\underline{P}|\gamma) | \gamma \in \Gamma\}$  be a family of distribution functions indexed by  $\gamma \in \Gamma$  which is continuous in  $\gamma$  and, for a fixed policy  $\underline{\pi}$ , let  $\mathcal{F}_{\underline{\pi}}$  be the corresponding family of marginal distribution functions,  $F_{\underline{\pi}}(P|\gamma)$ . Then  $\mathcal{F}_{\underline{\pi}}$  is a family of distributions continuous in  $\gamma$ .

Proof. In Theorem 2.4.3, let  $g(\underline{Q}) = 1$ ,  $\underline{Q} \in \mathcal{S}_{K,N}$  and, for fixed  $P(\underline{\pi}) \in \mathcal{S}_N$ , let

$$S = \left\{ \underline{Q} \mid \underline{Q} \in \mathcal{S}_{K,N}; q_{ij}^{\underline{Q}} \leq p_{ij}^{\underline{\pi}} \quad (i, j=1, \dots, N) \right\}. \quad (2.4.8)$$

Then equation (2.4.5) becomes

$$F_{\underline{\pi}}(P|\gamma) = \int_S dH(\underline{Q}|\gamma), \quad (2.4.9)$$

and  $F_{\underline{\pi}}(P|\gamma)$  is a continuous function of  $\gamma$  at  $P$  for any  $P \in \mathcal{S}_N$ .

Q.E.D.

Corollary 2.4.5 Let  $\mathcal{H} = \{H(\underline{P}|\gamma) | \gamma \in \Gamma\}$  be a family of

\* Cf., for example, Loève [23], pp. 180-182.



distributions indexed by  $\Psi$ . Suppose that, for every  $H(\underline{\varphi}|\Psi) \in \mathcal{H}$ , a corresponding density function  $h(\underline{\varphi}|\Psi)$  exists and that  $h(\underline{\varphi}|\Psi)$  is a continuous function of  $\Psi$  for every  $\underline{\varphi} \in \mathcal{S}_{K,N}$ . Then  $\mathcal{H}$  is a family of distributions continuous in  $\Psi$ .

*Proof.* The corollary follows immediately from a well-known theorem of integral calculus which states that, if  $h(\underline{Q}|\Psi)$  is continuous in  $\Psi$ , then

$$H(\underline{\varphi}|\Psi) = \int_S h(\underline{Q}|\Psi) d\underline{Q}$$

is a continuous function of  $\Psi$ , where, for a fixed  $\underline{\varphi} \in \mathcal{S}_{K,N}$ ,

$$S = \left\{ \underline{Q} \mid \underline{Q} \in \mathcal{S}_{K,N}; q_{ij}^k \leq p_{ij}^k, \quad (k=1, \dots, K_1; i, j=1, \dots, N) \right\}.$$

G.E.D.

It is clear from equations (2.2.1) and (2.2.2) that the matrix beta density function,  $f_{MB}^{(K,N)}(\underline{\varphi} | \underline{M})$ , is a continuous function of the  $K \times N$  matrix  $\underline{M}$ .



## CHAPTER 3

### ADAPTIVE CONTROL PROBLEMS

#### 3.1 Discounted Processes. Formulation.

Consider a Markov chain with alternatives in which the process, assumed to operate indefinitely, is sampled after each transition--that is, the decision-maker knows the state of the process after each transition.

Information about  $\underline{\Phi}$  is gathered in this manner and the decision-maker may alter the current policy at any time, as dictated by his state of knowledge about  $\widetilde{\underline{\Phi}}$ . Such a process is an adaptive control process.

It is assumed that any sampling costs are included in the transition reward matrix,  $\underline{Q} = [r_{ij}^k]$ . This implies that either the sampling costs are negligible when compared with the transition rewards or that the process is operated in such a manner that a sampling cost must be incurred after each transition. Models in which the decision-maker may choose to sample or not to sample will be considered in Chapter 5.

When future rewards are discounted to a present value we shall speak of a discounted adaptive control process. It is this class of problems which will be discussed, for the most part, in the present chapter. The interval between two consecutive transitions is assumed to be constant and can be taken as the time unit. Let  $\beta$  be the present value of a unit reward earned one unit of time in the future ( $0 \leq \beta < 1$ ). Since the present value of the maximum possible reward on the  $n$ th transition in the future decreases as  $\beta^n$ , it is clear that the total discounted reward earned over an infinite period under any sequence of policies is finite. A natural



criterion to use in choosing policies is, therefore, the expected total discounted reward over an infinite period and we shall define the discounted adaptive control problem to be the problem of selecting a sequence of policies so as to maximize this quantity.

In the present section the discounted adaptive control problem is formulated in terms of a set of simultaneous functional equations. It is shown in the following section that there exists a unique bounded set of continuous solutions to these equations. In Section 3.3 a method of successive approximations is described which converges monotonically and uniformly to this unique set of solutions and the question of policy convergence is considered. The concept of recursive computation is then introduced and a numerical example is presented. The chapter concludes with a discussion of the problems involved in treating undiscounted adaptive control processes in a Markov chain.

A specific form of the discounted adaptive control problem--the two-armed bandit problem--was treated by Bellman [7] in 1956, using dynamic programming and a beta prior distribution. The method was generalized by Bellman and Kalaba [8] and is summarized by Bellman in Chapter 16 of Adaptive Control Processes [6]. Bellman's method of solution is based upon the use of successive approximations.

Cazzolino [13] applied Bellman's formulation of the two-armed bandit problem to the case of a two-state Markov chain with two alternatives in each state, assuming a matrix beta prior distribution. He mapped decision regions in the parameter space of the prior distribution for the special case of one unknown transition probability vector. Cazzolino, Gonzalez-Zubieta, and Miller [14] have recently suggested various heuristic treatments of the discounted adaptive control problem, basing their results



on simulation studies. Freimer [18, 19] has obtained a solution of the discounted adaptive control problem in the case of quadratic cost functions by reducing the stochastic formulation to a deterministic one in terms of certainty equivalents.

The functional equations formulated in this chapter generalize the results of these authors and, in spirit, follow Bellman's derivation [6]. Our contribution to the treatment of this problem consists of the following:

- a. Proof of the existence of a unique bounded set of continuous solutions to these functional equations.
- b. Derivation of a method of successive approximations which converges monotonically and uniformly to this unique set of solutions.
- c. Introduction of recursive computation techniques for the numerical solution of the discounted adaptive control problem.

Let the prior distribution of  $\underline{\tilde{P}}$ ,  $H(\underline{\tilde{P}}|\Psi)$ , be a member of a family,  $\mathcal{H}$ , indexed by  $\Psi \in \Psi$ . The ordered pair,  $(i, \Psi)$ , where  $i = 1, \dots, N$  and  $\Psi \in \Psi$ , can be regarded as the generalized state of the system. Here,  $i$  is the physical state of the system and  $\Psi$  summarizes--or, more precisely, indexes--the decision-maker's state of knowledge about  $\underline{\tilde{P}}$ . Since the process is to be sampled consecutively, it must be assumed that  $\mathcal{H}$  is closed under the consecutive sampling rule in order that we may meaningfully refer to  $\Psi$  as indexing the decision-maker's state of knowledge as sampling progresses.

Let  $v_i(\Psi)$  denote the supremum of the expected discounted reward over an infinite period when the system starts from the generalized state,  $(i, \Psi)$ . If  $R = \max_{i,j,k} \{r_{i,j}^k\}$ , the discounted total reward under any sampling strategy is bounded by



$$\sum_{n=0}^{\infty} \beta^n R = \frac{R}{1-\beta} \quad (3.1.1)$$

and, therefore,  $v_i(\gamma)$  exists for  $i=1, \dots, N$  and all  $\gamma \in \Gamma$ . It will be shown at the conclusion of this section that  $v_i(\gamma)$  is attained under an optimal sampling strategy and, hence, can be regarded as the maximum expected discounted reward when the system starts from  $(i, \gamma)$ .

If, when in state  $(i, \gamma)$ , it is decided to choose the  $k$ th alternative and the system makes a transition to state  $j$ , the supremum of the posterior expected discounted reward is

$$r_{ij}^k + \beta v_j(T_{ij}^k(\gamma)). \quad (3.1.2)$$

The probability of the sample outcome  $j$ , unconditional with regard to the prior distribution of  $\underline{P}$ , given that the system is in state  $(i, \gamma)$  and that alternative  $k$  is in use, is

$$\tilde{p}_{ij}^k(\gamma) = \int_{\underline{P}_{k,N}} p_{ij}^k d\underline{H}(\underline{P}|\gamma), \quad (3.1.3)$$

the marginal prior expectation of  $\tilde{p}_{ij}^k$ . Let

$$\bar{q}_i^k(\gamma) = \sum_{j=1}^N \tilde{p}_{ij}^k(\gamma) r_{ij}^k \quad \begin{matrix} k=1, \dots, K_1 \\ i=1, \dots, N \\ \gamma \in \Gamma \end{matrix} \quad (3.1.4)$$

denote the mean one-step transition reward when the system is in state  $(i, \gamma)$  and alternative  $k$  is used. Then, regarding each  $v_i(\gamma)$  as a function of  $\gamma$  defined on  $\Gamma$  ( $i=1, \dots, N$ ), the supremum of the discounted expected reward when starting from  $(i, \gamma)$  must satisfy the following set of simultaneous functional equations,

$$v_i(\gamma) = \max_{1 \leq k \leq K_1} \left\{ \bar{q}_i^k(\gamma) + \beta \sum_{j=1}^N \tilde{p}_{ij}^k(\gamma) v_j(T_{ij}^k(\gamma)) \right\}. \quad (3.1.5)$$

$\begin{matrix} i=1, \dots, N \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{matrix}$



We now consider the existence of the maximum expected discounted reward over an infinite period. In order to do this, it is necessary to precisely define the notion of a sampling strategy for an adaptive control process.

Let the policies,  $\underline{\Sigma} \subset \epsilon\Sigma$ , be indexed by the integers 0 through  $J-1$ , where  $J$  is the number of elements in  $\Sigma$ . Thus,  $\Sigma = \{\underline{\sigma}_0, \underline{\sigma}_1, \dots, \underline{\sigma}_{J-1}\}$ . Suppose the system starts from the generalised state  $(i_0, \emptyset)$  and that alternative  $k$  has been selected in state  $i_0$ . We can, before the first transition occurs, decide which alternative to use in each state  $j$  for the second transition. This consists of the choice of a policy,  $\underline{\sigma}_{a_1}$ , and can be denoted by  $d_1(i_0, k) = a_1$ , a function with range  $\{0, 1, \dots, J-1\}$ .

In general, before any transitions have occurred, we can prescribe a policy to be used immediately after the  $n$ th transition ( $n=1, 2, \dots$ ).

Let  $x_{n-1} = (i_0, i_1, \dots, i_{n-1})$  be a possible sample history of the first  $n-1$  transitions and let  $s_{n-1} = (k, \underline{\sigma}_{a_1}, \dots, \underline{\sigma}_{a_{n-1}})$  be the sequence of policies under which the sample  $x_{n-1}$  occurred, together with  $\underline{\sigma}_{a_{n-1}}$ , the policy under which the  $n$ th transition will occur. The policy history,  $s_{n-1}$ , is determined by evaluating the decision functions  $d_1(i_0, k)$ ,  $d_2(x_1, s_1), \dots, d_{n-1}(x_{n-2}, s_{n-2})$  at  $i_0$ ,  $x_1 = (i_0, i_1), \dots, x_{n-2} = (i_0, i_1, \dots, i_{n-2})$ . Conditional on the Markov chain having arrived in state  $i_{n-1}$  with sample history  $x_{n-1}$  and policy history  $s_{n-1}$ , we may select an alternative for use in state  $j$  after the  $n$ th transition for each of the  $N$  states to which the  $n$ th transition may bring the system. This consists of the selection of a policy,  $\underline{\sigma}_{a_n} \in \Sigma$ , and is denoted by  $d_n(x_{n-1}, s_{n-1}) = a_n$ , a function with range  $\{0, \dots, J-1\}$ . Since there



are  $N^{(n-1)}$  different sample histories,  $x_{n-1}$ , which start from a fixed state  $i_0$ , it is necessary to specify  $N^{(n-1)}$  values of the nth-level decision function  $d_n(x_{n-1}, s_{n-1})$ . The specification of a complete set of decision functions,  $d_n(x_{n-1}, s_{n-1})$ , for  $n=1, 2, 3, \dots$  and for all possible sample histories, together with the choice of an initial alternative in state  $i_0$ , constitutes a sampling strategy,  $d$ . Let  $D_1$  denote the set of all possible sampling strategies when the system starts in state  $i$ .

In the following theorem it is shown that  $v_i(\gamma)$ , a least upper bound, is attained for some strategy  $d^* \in D_1$ . This is done by mapping the space of strategies  $D_1$  onto a compact subset of the real line and showing that the corresponding mapping of  $v_i(\gamma, d)$ , the expected discounted reward under strategy  $d$ , is continuous on this set.

**Theorem 3.1.1** Let  $v_i(\gamma, d)$  be the total expected discounted reward in a Markov chain with alternatives when the process starts from the generalized state  $(i, \gamma)$  and the strategy  $d$  is used. Let

$$v_i(\gamma) = \sup_{d \in D_1} \{v_i(\gamma, d)\}. \quad (3.1.6)$$

Then there is a strategy  $d^* \in D_1$  such that

$$v_i(\gamma) = v_i(\gamma, d^*). \quad (3.1.7)$$

**Proof.** For  $n=1, 2, 3, \dots$  let  $\xi_n = (i_0, i_1, \dots, i_n)$ , where  $i_a \in \{1, \dots, N\}$ ,  $a=0, 1, \dots, n$ . To each sequence  $\xi_n$  let there correspond the  $N$ -ary number

$$z(\xi_n) = \sum_{c=0}^n i_c N^{-(c+1)} = .i_0 i_1 \dots i_n. \quad (3.1.8)$$

For fixed  $n$  and  $i_0 = i$  we may then order the  $N^{(n-1)}$  different nth-level decision functions  $d_n(x_{n-1}, s_{n-1})$  as follows:  $d_n(x_{n-1}, s_{n-1}) < d_n(x'_{n-1}, s'_{n-1})$



if and only if  $v(z_{n-1}) < v(x'_{n-1})$ .

Consider a strategy  $d \in D_1$ . Let the value of the  $j$ th member of the  $\mu(n) = N^{(n-1)}$  decision functions  $d_n(x_{n-1}, s_{n-1})$  in  $d$  be denoted  $d_{nj}$  ( $n=1, 2, \dots$ ) and assume that the indexing is such that

$$d_{n1} < d_{n2} < \dots < d_{n\mu(n)}. \quad n=1, 2, \dots \quad (3.1.9)$$

This can be done since the  $\mu(n)$  possible sample histories  $x_{n-1}$  which lead to the  $n$ th level decision functions are all distinct. The strategy  $d$  can then be displayed as the ordered pair

$$d = (k, \delta) \quad (3.1.10)$$

where  $k$  is the initial alternative selected in state 1 and  $\delta$  is the sequence

$$\begin{aligned} \delta &= \{d_{11}, d_{21}, d_{31}, \dots\} \\ &= \{\delta_1, \delta_2, \delta_3, \dots\}. \end{aligned} \quad (3.1.11)$$

Letting  $\Delta$  denote the set of all possible sequences  $\delta$ , we have

$$v_1(\gamma) = \max_{1 \leq k \leq K_1} \sup_{\delta \in \Delta} \{v_1(\gamma; k, \delta)\}, \quad (3.1.12)$$

where, if  $d = (k, \delta)$ ,  $v_1(\gamma; k, \delta) = v_1(\gamma, d)$ .

To each  $\delta \in \Delta$  let there correspond the  $J$ -ary number

$$y(\delta) = \sum_{c=1}^{\infty} \delta_c J^{-c} = .\delta_1 \delta_2 \delta_3 \dots, \quad (3.1.13)$$

where  $\delta_c \in \{0, 1, \dots, J-1\}$  for  $c=1, 2, \dots$ . If  $\delta = \{0, 0, 0, \dots\}$  then  $y(\delta) = 0$  and if  $\delta = \{J-1, J-1, J-1, \dots\}$  then

$$y(\delta) = (J-1) \sum_{c=1}^{\infty} J^{-c} = \frac{(J-1)J^{-1}}{1-J^{-1}} = 1. \quad (3.1.14)$$

Thus, for any  $\delta \in \Delta$ ,  $0 \leq y(\delta) \leq 1$ . Moreover, it is easily seen that the mapping (3.1.13) is a one-to-one mapping of the set  $\Delta$  onto the closed



interval  $[0,1]$ . Let  $w_2^k(\psi, y)$  be a function defined on  $[0,1]$  by the relation

$$w_2^k(\psi, y(\delta)) = v_1(\psi; k, \delta), \quad \begin{array}{l} k=1, \dots, K \\ i=1, \dots, N^1 \\ \delta \in \Delta; \quad \psi \in \Psi \end{array} \quad (3.1.15)$$

Then (3.1.12) can be written

$$v_1(\psi) = \max_{1 \leq k \leq K_1} \sup_{0 \leq y \leq 1} \{ w_2^k(\psi, y) \}. \quad (3.1.16)$$

We now show that, for fixed  $k$ ,  $w_2^k(\psi, y)$  is continuous in  $y$ . Let

$$R^* = \max_{i,j,k} \{ |r_{ij}^k| \}. \quad (3.1.17a)$$

and

$$r^* = \min_{i,j,k} \{ |r_{ij}^k| \}. \quad (3.1.17b)$$

Let  $\epsilon > 0$  be given and choose a positive integer  $v$  such that

$$\beta^{v+1} \frac{\frac{\epsilon}{R^*}}{1-\beta} < \epsilon. \quad (3.1.18)$$

For a fixed  $y \in [0,1]$  let  $y^*$  be any number such that  $0 \leq y^* \leq 1$  and

$|y - y^*| < \beta^{-v}$ . Then, if  $y = y(\delta)$  and  $y^* = y^*(\delta^*)$ , we have

$$\delta_a = \delta_a^* \quad \text{or } 1, 2, \dots, v \quad (3.1.19)$$

and

$$\begin{aligned} |w_2^k(\psi, y) - w_2^k(\psi, y^*)| &\leq \sum_{a=v+1}^{\infty} \beta^a (R^* - r^*) \\ &= \beta^{v+1} \frac{R^* - r^*}{1-\beta} < \epsilon \end{aligned} \quad (3.1.20)$$

Thus,  $w_2^k(\psi, y)$  is a continuous function of  $y$  on the compact set  $[0,1]$  and, for each  $k$ , there exists a  $y_k^* \in [0,1]$  such that

$$w_2^k(\psi, y_k^*) = \sup_{0 \leq y \leq 1} \{ w_2^k(\psi, y) \}. \quad (3.1.21)$$

Letting  $\delta^*(k)$  denote the inverse image of  $y_k^* = y_k^*(\delta^*)$ ,

$$v_1(\psi) = \max_{1 \leq k \leq K_1} \{ v_1(\psi; k, \delta^*(k)) \} \quad (3.1.22)$$

and there exists a strategy,  $d^* = (k^*, \delta^*(k^*))$ , such that



Q.E.D.

$$v_1(\psi) = v_1(\psi, d^*). \quad (3.1.23)$$

### 3.2 Existence and Uniqueness of $v_1(\psi)$ .

In order to show the existence of a unique bounded set of continuous solutions to the functional equations (3.1.5) we shall make use of the method of successive approximations. Let the functions  $v_1(n, \psi)$  be defined recursively as follows,

$$v_1(n+1, \psi) = \max_{1 \leq k \leq K_1} \left\{ q_1^k(\psi) + \beta \sum_{j=1}^N p_{1j}^k(\psi) v_j(n, T_{1j}^k(\psi)) \right\},$$

(3.2.1)

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

and

$$v_1(0, \psi) = v_1(\psi), \quad i=1, \dots, N \quad (3.2.2)$$

$\psi \in \Psi$

where  $v_1(\psi)$  is a set of bounded terminal functions.

It will be convenient to introduce some additional notation. Let

$$s_1^k(v, n, \psi) = q_1^k(\psi) + \beta \sum_{j=1}^N p_{1j}^k(\psi) v_j(n, T_{1j}^k(\psi)), \quad k=1, \dots, K_1$$

(3.2.3)

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$

and

$$s_1^k(v, \infty, \psi) = q_1^k(\psi) + \beta \sum_{j=1}^N p_{1j}^k(\psi) v_j(T_{1j}^k(\psi)).$$

(3.2.4)

$k=1, \dots, K_1$   
 $i=1, \dots, N$   
 $\psi \in \Psi$

Equation (3.2.1) may then be written

$$v_1(n+1, \psi) = \max_{1 \leq k \leq K_1} \{ s_1^k(v, n, \psi) \}, \quad i=1, \dots, N \quad (3.2.5)$$

$n=0, 1, 2, \dots$   
 $\psi \in \Psi$

and, similarly, equation (3.1.5) becomes

$$v_1(\psi) = \max_{1 \leq k \leq K_1} \{ s_1^k(v, \infty, \psi) \}. \quad i=1, \dots, N \quad (3.2.6)$$

$\psi \in \Psi$



The first result is a lemma which will be used in subsequent proofs.

**Lemma 3.2.1** If  $V$  is a bound for the terminal functions  $v_1(\tau)$ ,

$$|v_1(\tau)| \leq V, \quad \begin{matrix} i=1, \dots, N \\ \tau \in \mathbb{P} \end{matrix} \quad (3.2.7)$$

and if

$$R^* = \max_{1, j, k} \left\{ |v_{1j}^k| \right\}, \quad (3.2.8)$$

then the functions  $v_1(n, \tau)$  are bounded,

$$|v_1(n, \tau)| \leq \beta^n V + \frac{1 - \beta^n}{1 - \beta} R^*. \quad (3.2.9)$$

$\begin{matrix} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \tau \in \mathbb{P} \\ 0 \leq \beta < 1 \end{matrix}$

**Proof.** The proof is inductive. Equation (3.2.9) obviously holds for  $n = 0$ . Assume it holds for  $n$ . Then, if  $k = a$  maximizes the right side of (3.2.1),

$$\begin{aligned} |v_1(n+1, \tau)| &\leq |v_1^a(\tau)| + \beta \sum_{j=1}^N p_{1j}^a(\tau) |v_j(n, v_{1j}^a(\tau))| \\ &\leq R^* + \beta [ \beta^n V + \frac{1 - \beta^n}{1 - \beta} R^* ] \\ &= \beta^{n+1} V + \frac{1 - \beta^{n+1}}{1 - \beta} R^*. \end{aligned} \quad (3.2.10)$$

Q.E.D.

**Theorem 3.2.2** If the set of functions  $\{v_1(n, \tau)\}$  is defined by equations (3.2.1) and (3.2.2), then the limits

$$\lim_{n \rightarrow \infty} v_1(n, \tau) = v_1(\tau) \quad \begin{matrix} i=1, \dots, N \\ \tau \in \mathbb{P} \end{matrix} \quad (3.2.11)$$

exist and  $\{v_1(\tau)\}$  is a set of solutions to equation (3.1.5). Moreover, the convergence is uniform in  $\tau$ .

**Proof.** It will be established inductively that, for arbitrary positive integers,  $n$  and  $n'$ ,



$$\left| v_1(n, \gamma) - v_1(m, \gamma) \right| \leq (\beta^n + \beta^m)V + \frac{|\beta^n - \beta^m|}{1-\beta} R^*, \quad (3.2.12)$$

$i=1, \dots, N$   
 $n, m=0, 1, 2, \dots$   
 $\gamma \in \mathbb{P}$

where  $V$  is a bound on the terminal functions. Since  $0 \leq \beta < 1$ , it then follows by the Cauchy criterion that  $\lim_{n \rightarrow \infty} v_1(n, \gamma)$  exists for  $i=1, \dots, N$ . That the limiting functions satisfy (3.1.5) follows by allowing  $n$  to go to  $\infty$  in equation (3.2.1). Uniform convergence of the sequence of functions  $\{v_1(n, \cdot)\}$  follows by noting that the bound (3.2.12) is independent of  $\gamma$ .

To establish (3.2.12) we proceed as follows. Using the formulation (3.2.5), for any fixed  $\gamma \in \mathbb{P}$  let

$$v_1(n, \gamma) = s_1^a(v, n-1, \gamma) = \max_{1 \leq k \leq K_1} \{s_1^k(v, n-1, \gamma)\}$$

$$v_1^*(n, \gamma) = s_1^b(v, n-1, \gamma) = \max_{1 \leq k \leq K_1} \{s_1^k(v, n-1, \gamma)\}.$$

Then

$$v_1(n, \gamma) - v_1^*(n, \gamma) = s_1^a(v, n-1, \gamma) - s_1^b(v, n-1, \gamma)$$

$$\leq s_1^a(v, n-1, \gamma) - s_1^{a^*}(v, n-1, \gamma)$$

and, similarly,

$$v_1(n, \gamma) - v_1^*(n, \gamma) \geq s_1^b(v, n-1, \gamma) - s_1^{b^*}(v, n-1, \gamma).$$

Let  $k^*$  index the maximum of  $|s_1^a(v, n-1, \gamma) - s_1^{a^*}(v, n-1, \gamma)|$  and

$$|s_1^b(v, n-1, \gamma) - s_1^{b^*}(v, n-1, \gamma)|. \quad \text{Then}$$

$$\left| v_1(n, \gamma) - v_1^*(n, \gamma) \right| \leq \left| s_1^{k^*}(v, n-1, \gamma) - s_1^{k^*}(v, n-1, \gamma) \right|$$

$$\leq \beta \sum_{j=1}^N p_{1j}^{k^*}(\gamma) \left| v_j(n-1, T_{1j}^{k^*}(\gamma)) - v_j(n-1, T_{1j}^{k^*}(\gamma)) \right|. \quad (3.2.13)$$

$i=1, \dots, N$   
 $n, m=1, 2, \dots$   
 $\gamma \in \mathbb{P}$



Assuming that  $n > m$ , Lemma 3.2.1 implies the inequality

$$|v_1(n-m, \psi) - v_1(0, \psi)| \leq (1 + \beta^{n-m}) v + \frac{1 - \beta^{n-m}}{1 - \beta} R^*. \quad (3.2.14)$$

An inductive argument, using (3.2.13), shows that

$$|v_1(n, \psi) - v_1(m, \psi)| \leq (\beta^m + \beta^n) v + \frac{\beta^m - \beta^n}{1 - \beta} R^*. \quad (3.2.15)$$

A similar argument in the case  $n < m$  yields (3.2.12). Q.E.D.

**Theorem 3.2.3** There exists a unique set of bounded functions

$\{v_i(\psi)\}$  which satisfies the set of equations

$$v_i(\psi) = \max_{1 \leq k \leq K_i} \left\{ \bar{q}_i^k(\psi) + \beta \sum_{j=1}^N \bar{p}_{ij}^k(\psi) v_j(t_{ij}^k(\psi)) \right\}. \quad (3.2.16)$$

$i=1, \dots, N$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

**Proof.** Theorem 3.2.2 established the existence of at least one set of functions  $\{v_i(\psi)\}$  which satisfy (3.2.16). Lemma 3.2.1 implies that this set of functions is bounded:

$$|v_i(\psi)| = \lim_{n \rightarrow \infty} |v_i(n, \psi)| \leq \frac{R^*}{1-\beta}. \quad i=1, \dots, N \quad \psi \in \Psi \quad (3.2.17)$$

To establish uniqueness, assume that there exist two sets of bounded functions,  $\{v_i(\psi)\}$  and  $\{w_i(\psi)\}$ , which satisfy (3.2.16). For any  $i$  and an arbitrary  $\psi \in \Psi$ , let

$$\begin{aligned} v_i^a(\psi) &= S_i^a(v, \infty, \psi) \\ w_i^b(\psi) &= S_i^b(w, \infty, \psi). \end{aligned}$$

Then, arguing as in the proof of Theorem 3.2.2,

$$S_i^b(v, \infty, \psi) - S_i^b(w, \infty, \psi) \leq v_i(\psi) - w_i(\psi) \leq S_i^a(v, \infty, \psi) - S_i^a(w, \infty, \psi) \quad (3.2.18)$$

Letting  $k^*$  index the maximum of  $|S_i^a(v, \infty, \psi) - S_i^a(w, \infty, \psi)|$  and

$$|S_i^b(v, \infty, \psi) - S_i^b(w, \infty, \psi)|,$$



$$\left| v_1(\psi) - w_1(\psi) \right| \leq \beta \sum_{j=1}^N \bar{p}_{1,j}^{k^*}(\psi) \left| v_1(T_{1,j}^{k^*}(\psi)) - w_1(T_{1,j}^{k^*}(\psi)) \right|. \\ \begin{matrix} i=1, \dots, N \\ \psi \in \mathbb{F} \end{matrix} \quad (3.2.19)$$

Since  $v_1(\psi)$  and  $w_1(\psi)$  are both bounded functions of  $\psi$ , there exists a number,  $M > 0$ , such that

$$\left| v_1(\psi) - w_1(\psi) \right| \leq M. \quad \begin{matrix} i=1, \dots, N \\ \psi \in \mathbb{F} \end{matrix} \quad (3.2.20)$$

Repeated application of (3.2.19) then yields the inequality

$$\left| v_1(\psi) - w_1(\psi) \right| \leq \beta^n M. \quad \begin{matrix} n=0, 1, 2, \dots \\ i=1, \dots, N \\ \psi \in \mathbb{F} \end{matrix} \quad (3.2.21)$$

Since  $0 \leq \beta < 1$ , it then follows that

$$v_1(\psi) = w_1(\psi). \quad \begin{matrix} i=1, \dots, N \\ \psi \in \mathbb{F} \end{matrix} \quad (3.2.22)$$

Q.E.D.

**Theorem 3.2.4** If  $\{v_1(\psi)\}$  is the unique bounded set of functions which satisfy equation (3.1.5), and if  $\bar{p}_{1,j}^k(\psi)$  is a continuous function of  $\psi$  ( $k=1, \dots, K_1$ ;  $i, j=1, \dots, N$ ), then  $v_1(\psi)$  is a continuous function of  $\psi$  ( $i=1, \dots, N$ ).

**Proof.** Consider the functions  $v_1(n, \psi)$  defined by equations (3.2.1) and (3.2.2). Choosing a set of terminal functions  $\{V_1(\psi)\}$  each member of which is continuous on  $\mathbb{F}$ , it follows inductively that  $v_1(n, \psi)$  is continuous ( $i=1, \dots, N$ ;  $n=0, 1, 2, \dots$ ). By Theorem 3.2.2  $\{v_1(n, \psi)\} \rightarrow v_1(\psi)$  uniformly and, therefore,  $v_1(\psi)$  is continuous.

Q.E.D.

### 3.3 Solution by Successive Approximations.

The functions  $v_1(n, \psi)$  defined by equations (3.2.1) and (3.2.2) can



be used as successive approximations in the numerical calculation of  $v_i(\gamma)$  at some fixed generalized state,  $(i, \gamma)$ . In this section we derive conditions under which the sequence of functions  $\{v_i(n, \gamma)\}$  converges monotonically and find a bound for the error of the nth approximant  $e_i(n, \gamma) = v_i(\gamma) - v_i(n, \gamma)$ . The section concludes with a proof that the optimal sampling strategy of the n-step problem defined by (3.2.1) and (3.2.2) converges to an optimal sampling strategy for the infinite horizon problem defined by (3.1.5).

**Theorem 3.3.1** Let the terminal functions of equation (3.2.2) be constants,

$$v_i(\gamma) = V_i. \quad \begin{matrix} i=1, \dots, N \\ \gamma \in \Xi \end{matrix} \quad (3.3.1)$$

Let

$$v^* = \min_i \{v_i\} \quad v^* = \max_i \{v_i\} \quad (3.3.2)$$

and

$$r = \min_{i,j,k} \{r_{ij}^k\} \quad R = \max_{i,j,k} \{r_{ij}^k\}. \quad (3.3.3)$$

If

$$v^* - \beta v^* \leq r \quad (3.3.4)$$

then, for  $i=1, \dots, N$ , the functions  $v_i(n, \gamma)$  defined by equations (3.2.1) and (3.2.2) form a monotone increasing sequence which converges uniformly to  $v_i(\gamma)$ , the unique bounded solution of equation (3.1.5).

Similarly, if

$$v^* - \beta v^* \geq R, \quad (3.3.5)$$

then the functions  $v_i(n, \gamma)$  form a monotone decreasing sequence which converges uniformly to  $v_i(\gamma)$ .

Proof. We will first show inductively that, if (3.3.4) holds, the sequences  $\{v_i(n, \gamma)\}$  are monotone increasing for each  $i=1, \dots, N$



and each  $\psi \in \Psi$ . Uniform convergence of  $\{v_1(n, \psi)\}$  to  $v_1(\psi)$  has already been demonstrated. If equation (3.3.4) is satisfied, then

$$\begin{aligned} v_1(1, \psi) - v_1(0, \psi) &= \max_{1 \leq k \leq K_1} \left\{ \bar{q}_2^k(\psi) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\psi) v_j \right\} - v_1 \\ &\geq r + \beta v^* - v^* \geq 0. \end{aligned} \quad \begin{matrix} i=1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (3.3.6)$$

Assume that  $v_1(n, \psi) = v_1(n-1, \psi) \geq 0$  for  $i=1, \dots, N$  and  $\psi \in \Psi$ . Then,

if

$$v_1(n+1, \psi) = S_1^\delta(v, n, \psi)$$

and

$$v_1(n, \psi) = S_1^\delta(v, n-1, \psi),$$

the inductive hypothesis implies that

$$\begin{aligned} v_1(n+1, \psi) - v_1(n, \psi) &= S_1^\delta(v, n, \psi) - S_1^\delta(v, n-1, \psi) \\ &\geq S_1^\delta(v, n, \psi) - S_1^\delta(v, n-1, \psi) \\ &= \beta \sum_{j=1}^N \bar{p}_{1j}^\delta(\psi) [v_j(n, T_{1j}^\delta(\psi)) - v_j(n-1, T_{1j}^\delta(\psi))] \\ &\geq 0, \end{aligned} \quad \begin{matrix} i=1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (3.3.7)$$

proving the induction. If (3.3.5) holds,

$$v_1(1, \psi) - v_1(0, \psi) \leq r + \beta v^* - v^* \leq 0. \quad \begin{matrix} i=1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (3.3.8)$$

That  $v_1(n+1, \psi) \leq v_1(n, \psi)$  is then easily established by induction in a manner similar to that displayed in equation (3.3.7). Q.E.D.

Let the error of the  $n$ th approximant be defined as

$$e_1(n, \psi) = v_1(\psi) - v_1(n, \psi). \quad \begin{matrix} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \psi \in \Psi \end{matrix} \quad (3.3.9)$$



If  $\{v_1(n, \psi)\}$  is a sequence of functions which converges monotonically to  $v_1(\psi)$  then  $\{e_1(n, \psi)\}$  is a sequence of functions which converges monotonically to zero. In this case, if  $\epsilon > 0$  is an error-bound which is acceptable to the decision-maker and if  $n^*$  is the smallest positive integer such that

$$|e_1(n, \psi)| \leq \epsilon, \quad (3.3.10)$$

then  $v_1(n^*, \psi)$  is an acceptable approximation to  $v_1(\psi)$  and the sampling strategy resulting in  $v_1(n^*, \psi)$  is an acceptable approximation to the first  $n^*$  levels of an optimal sampling strategy.

It is not necessary to require that the successive approximants  $v_1(n, \psi)$  converge monotonically to  $v_1(\psi)$  and, in fact, a non-monotonic sequence  $\{v_1(n, \psi)\}$  may converge more rapidly than a monotonic sequence. Theorem 3.3.3 provides a bound for  $e_1(n, \psi)$  assuming nothing about monotonicity. A lemma is first required.

Lemma 3.3.2 Let  $r$  and  $R$  be defined by equation (3.3.3). Then  $v_1(\psi)$ , the solution of (3.1.5), has the bounds

$$\frac{r}{1-\beta} \leq v_1(\psi) \leq \frac{R}{1-\beta}. \quad \begin{matrix} i=1, \dots, N \\ \psi \in \Psi \\ 0 \leq \beta < 1 \end{matrix} \quad (3.3.11)$$

Proof. The mean reward per transition under any policy has the bounds

$$r \leq \bar{q}_1^k(\psi) \leq R. \quad \begin{matrix} k=1, \dots, K_1 \\ i=1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (3.3.12)$$

Since the expected discounted reward over an infinite period under any strategy is the sum of the expected rewards at each transition, the reward of the  $n$ th transition being weighted by  $\beta^n$ , the maximum total reward over all strategies has the bounds



-54-

$$r \sum_{n=0}^{\infty} \beta^n \leq v_1(\psi) \leq R \sum_{n=0}^{\infty} \beta^n, \quad i=1, \dots, N \quad (\text{3.3.13})$$

$\psi \in \Psi$

from which (3.3.11) follows. Q.E.D.

**Theorem 3.3.3** Let  $\{v_i(n, \psi)\}$  be a sequence of successive approximations defined by equations (3.2.1) and (3.2.2), with constant terminal reward functions,

$$v_i(\psi) = v_i. \quad i=1, \dots, N \quad (\text{3.3.14})$$

Let  $v^*$ ,  $V^*$ ,  $r$ , and  $R$  be defined by (3.3.2) and (3.3.3). Then the error of the  $n$ th approximant has the bound

$$|e_i(n, \psi)| \leq \beta^n [\max \left\{ \frac{R}{1-\beta} - v^*, V^* - \frac{R}{1-\beta} \right\}]. \quad (\text{3.3.15})$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

**Proof.** The proof is inductive. For  $n=0$ ,

$$e_i(0, \psi) = v_i(\psi) - v_i.$$

Suppose  $v_i(\psi) \geq v_i$ . Then, by Lemma 3.3.2,

$$|e_i(0, \psi)| = v_i(\psi) - v_i \leq \frac{R}{1-\beta} - v^*. \quad i=1, \dots, N \quad (\text{3.3.16})$$

$\psi \in \Psi$

If, on the other hand,  $v_i(\psi) < v_i$ , Lemma 3.3.2 implies

$$|e_i(0, \psi)| = v_i - v_i(\psi) \leq V^* - \frac{R}{1-\beta}. \quad i=1, \dots, N \quad (\text{3.3.17})$$

$\psi \in \Psi$

Therefore, in either case,

$$|e_i(0, \psi)| \leq \max \left\{ \frac{R}{1-\beta} - v^*, V^* - \frac{R}{1-\beta} \right\}. \quad i=1, \dots, N \quad (\text{3.3.18})$$

$\psi \in \Psi$

It is to be noted that at least one of the two terms,  $\frac{R}{1-\beta} - v^*$  and  $V^* - \frac{R}{1-\beta}$ , is non-negative, for, if not, we have the contradiction



$$v^* < \frac{r}{1-\beta} \leq \frac{R}{1-\beta} < v^* \leq V^*.$$

Having established that equation (3.3.15) is valid for  $n=0$ , assume it holds for  $n$ . Let

$$\begin{aligned} v_1(\tau) &= S_1^a(v, \infty, \tau) \\ v_1(n+1, \tau) &= S_1^a(v, n, \tau). \end{aligned}$$

Then

$$S_1^a(v, \infty, \tau) = S_1^a(v, n, \tau) \leq e_1(n+1, \tau) \leq S_1^a(v, \infty, \tau) - S_1^a(v, n, \tau). \quad (3.3.19)$$

Let  $k^*$  index the maximum of  $|S_1^a(v, \infty, \tau) - S_1^a(v, n, \tau)|$  and

$$|S_1^a(v, \infty, \tau) - S_1^a(v, n, \tau)|. \text{ Then}$$

$$\begin{aligned} |e_1(n+1, \tau)| &\leq |S_1^{k^*}(v, \infty, \tau) - S_1^{k^*}(v, n, \tau)| \\ &\leq \beta \sum_{j=1}^N p_{1j}^{k^*}(\tau) |e_j(n, \tau_{1j}^{k^*}(\tau))| \\ &\leq \beta^{n+1} [\max \left\{ \frac{R}{1-\beta} - v^*, V^* - \frac{r}{1-\beta} \right\}]. \end{aligned} \quad (3.3.20)$$

Q.E.D.

Corollary 3.3.4 Let the terminal functions  $v_i(\tau)$  be constants,

$$v_i(\tau) = v_i, \quad i=1, \dots, N \quad (3.3.21) \quad \tau \in \mathbb{E}$$

Then, if

$$V^* - \beta v^* \leq r, \quad (3.3.22)$$

the error of the  $n$ th approximant has the bounds

$$0 \leq e_1(n, \tau) \leq \beta^n \left( \frac{R}{1-\beta} - v^* \right). \quad (3.3.23)$$

$i=1, \dots, N$

$n=0, 1, 2, \dots$

$\tau \in \mathbb{E}$

$0 \leq \beta < 1$

Similarly, if



$$v^* = \beta V^* \geq R, \quad (3.3.24)$$

$e_1(n, \gamma)$  has the bounds

$$\beta^n \left( \frac{R}{1-\beta} - v^* \right) \leq e_1(n, \gamma) \leq 0. \quad \begin{array}{l} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{array} \quad (3.3.25)$$

**Proof.** If equation (3.3.22) is satisfied Theorem 3.3.1 implies that  $e_1(n, \gamma) \geq 0$ . Moreover, (3.3.22) implies that  $v^*(1-\beta) \leq R$  and  $V^*(1-\beta) \leq R$ , hence, that  $\frac{R}{1-\beta} - v^* \geq v^* - \frac{R}{1-\beta}$ . Equation (3.3.15) then yields the upper inequality of (3.3.23). Similarly, if equation (3.3.24) is satisfied, then  $V^*(1-\beta) \geq R$  and  $v^*(1-\beta) \geq R$ , hence  $V^* - \frac{R}{1-\beta} \geq \frac{R}{1-\beta} - v^*$ . The bounds of (3.3.25) then follow from Theorem 3.3.1 and equation (3.3.15). Q.E.D.

Let there correspond to the  $n$ th approximant  $v_1(n, \gamma)$ , defined by (3.2.1) and (3.2.2), the  $n$ -step optimal sampling strategy  $d^*(n)$ . At least one such optimal strategy exists since there are a finite number of different sampling strategies for the  $n$ -step problem; there may be more than one  $n$ -step optimal strategy. The next theorem demonstrates that, as  $n \rightarrow \infty$ , any  $n$ -step optimal strategy converges to an optimal sampling strategy for the adaptive control model of equation (3.1.5). We must first precisely define what is meant by convergence of a sampling strategy.

Let the generalized state  $(i, \gamma)$  be fixed. To every  $n$ -step sampling strategy  $d(n)$  there corresponds an ordered pair  $(k_n, y_n)$ , where  $k_n \in \{1, \dots, K_i\}$  and  $y_n \in [0, 1]$  is defined by equation (3.1.13) with  $s_c = 0$  for  $c > \sum_{k=1}^{K_i} n^{(k-1)}$ . Let  $d^* = (k^*, y^*)$  be a sampling strategy for the infinite horizon model of (3.1.5). Then we say  $\lim_{n \rightarrow \infty} d(n) = d^*$



i.e., given  $\epsilon > 0$ , there is a positive integer  $v$  such that, for all  $n > v$ ,  $k_n = k^*$  and  $|y_n - y^*| < \epsilon$ . This definition implies that, given an arbitrarily large positive integer  $\mu$ , there exists an integer  $v$  such that, for all  $n > v$ , the values of the decision functions on the first  $\mu$  levels of  $d(n)$  are equal to the values of the decision functions on the first  $\mu$  levels of  $d^*$ .

**Theorem 3.3.5** Let the generalized state  $(i, \tau)$  be fixed and let  $\Delta \subset D_i$  denote the set of optimal sampling strategies for the adaptive control problem of equation (3.1.5). If  $d^*(n)$  is an  $n$ -step optimal sampling strategy for the problem defined by (3.2.1) and (3.2.2), then

$$\lim_{n \rightarrow \infty} d^*(n) = d \quad (3.3.26)$$

exists and  $d \in \Delta$ .

**Proof.** Let  $k_n$  denote the initial alternative selected in the  $n$ -step optimal sampling strategy  $d^*(n)$  and let  $X_1$  be the set of alternatives in state  $i$  which are initial selections for an optimal strategy in  $\Delta$ . We first show that  $\lim_{n \rightarrow \infty} k_n = k \in X_1$ .

Using the notation of equations (3.2.3) and (3.2.4),

$$v_1(n, \tau) = s_1^{k_n}(v, n-1, \tau), \quad (3.3.27)$$

and, for any  $k \in X_1$  and  $n \in \mathbb{N}$ ,

$$v_1(\tau) = s_1^k(v, \infty, \tau) > s_1^n(v, \infty, \tau). \quad (3.3.28)$$

Assume that  $\{k_n\}$  does not converge to a member of  $X_1$ . Then there exists a subsequence  $\{k_{n_\nu}\}$  such that

$$k_{n_\nu} \notin X_1. \quad \forall i, 2, \dots \quad (3.3.29)$$

Let  $\epsilon$  be chosen such that

$$0 < \epsilon < \min_{a \in X_1} \left\{ |v_1(\tau) - s_1^a(v, \infty, \tau)| \right\}. \quad (3.3.30)$$



Since  $\lim_{v \rightarrow \infty} v_j(n_v, \gamma) = v_j(\gamma)$ , we have, by (3.3.27),

$$\left| v_j(\gamma) - s_1^{k_{n_v}}(v, n_v-1, \gamma) \right| < \frac{\epsilon}{2} \quad (3.3.31)$$

for all  $v$  sufficiently large. But, using (3.2.3) and (3.2.4),

$$\lim_{n \rightarrow \infty} s_1^k(v, n-1, \gamma) = s_1^k(v, \infty, \gamma) \quad k=1, \dots, K_j \quad (3.3.32)$$

and there exists an integer  $v$  such that

$$\left| s_1^{k_{n_v}}(v, n_v-1, \gamma) - s_1^{k_{n_v}}(v, \infty, \gamma) \right| < \frac{\epsilon}{2}. \quad (3.3.33)$$

Thus, combining (3.3.31) and (3.3.33), there exists an integer  $v$  such

that  $k_{n_v} \notin X_j$  and

$$\left| v_j(\gamma) - s_1^{k_{n_v}}(v, \infty, \gamma) \right| < \epsilon, \quad (3.3.34)$$

contradicting (3.3.30). It follows that  $\lim_{n \rightarrow \infty} k_n = k$  exists and that  $k \in X_j$ .

Given a positive integer  $\mu$ , the same proof can be applied to each selection of an alternative in the first  $\mu$  levels of the sampling strategy  $d^\mu(n)$ . Since, having fixed  $\mu$ , there are a finite number of such alternatives, there exists a positive integer  $v$  such that for all  $n \geq v$ , the decision functions in the first  $\mu$  levels of  $d^\mu(n)$  have the same values as the corresponding decision functions in some strategy  $d \in \Delta$ . Q.E.D.

### 3.4 Recursive Computation.

Equations (3.2.1) and (3.2.2) are typical of a class of recursive equations which appear throughout this report. It is to be noted that, although these equations resemble a classical iterative formula for successive approximations,  $v_j(n, \gamma)$  is computed, not in terms of  $v_j(n-1, \gamma)$ , but in terms of  $v_j(n-1, T_{1j}^k(\gamma))$ . Computation of  $v_j(n, \gamma)$  for a specific value of  $(j, \gamma)$  involves the evaluation of between



$(k_1)_j^{(n+1)}$  and  $(k_2)_j^{(n+1)}$  terminal values  $V_j(\gamma^*)$ , where  $k_1 = \min_i \{ k_i \}$  and  $k_2 = \max_i \{ k_i \}$ .

One way to compute  $v_j(n, \gamma)$  is to start by evaluating and storing all required values of  $V_j(\gamma) = v_j(0, \gamma)$ , then to compute and store all required values of  $v_j(1, \gamma)$ , using the results of the previous computation of  $v_j(0, \gamma)$ . In general, for  $v = 1, 2, \dots, n-1$ ,  $v_j(v, \gamma)$  is computed in terms of a grid of values of  $v_j(v+1, \gamma)$  and is stored for use at the next stage in computing  $v_j(v+1, \gamma)$ .

Since the number of terminal values  $V_j(\gamma)$  which are needed grows exponentially with  $n$ , it is clear that considerable storage capacity is required. For even moderately large  $n$ , tape or disc storage must be used. Moreover, a fairly complex indexing routine must be programmed in order to utilize core memory efficiently.

An alternative approach is to evaluate  $v_j(n, \gamma)$  recursively. Using this method, computation starts with the  $n$ th level rather than the zero-th level. In general, the routine, at the  $(v+1)$ th level of computation, starts to evaluate  $v_j(v+1, \gamma)$  for some pair  $(j, \gamma)$ . This level of computation is suspended when a value at the  $v$ th level,  $v_j(v, \gamma^*)$ , is required. Certain key portions of the  $(v+1)$ th level of computation are stored on a push-down list and the routine then calls itself, entering the  $v$ th level of computation to evaluate  $v_j(v, \gamma^*)$ . Recursion is halted at the zero-th level when  $V_j(\gamma)$  is computed. The results of lower level computations are then fed back, in succession, to higher levels. Having obtained the value of  $v_j(v, \gamma^*)$  in this manner, the  $(v+1)$ th level of computation reclaims its partially completed calculations from the push-down list and completes them. This succession of events continues until  $v_j(n, \gamma)$  is evaluated.



The storage requirements for recursive calculation of  $v_i(n, \psi)$  consist only of the space needed for storage of intermediate computations on the push-down list and, therefore, increase linearly with  $n$ . Thus, the recursive method has the advantage of requiring considerably less storage than the first method described. Since specific values of  $v_j(v, \psi)$ , for  $v = 0, 1, \dots, n-1$ , may have to be recalculated many times in the recursive method, we are essentially trading running time for storage. It should be noted, however, that if the first method requires tape handling, the recursive method may reduce overall running time.

The general theory of recursive computation is described by McCarthy [29]. Programming languages of the ALGOL family [32] are capable of recursive computations as are most list processing languages; it is possible to do recursive programming in FORTRAN II [4]. The recursive programs which were written for this report used the MAD language [3].

Utilizing the recursive method, a program was written to evaluate equations (3.2.1) and (3.2.2) for specific pairs  $(i, \psi)$  when  $\tilde{P}$  has the matrix beta distribution. This program is contained in Appendix B. Some numerical results obtained from the program are presented in the next section.

### 3.5 Numerical Example.

Consider a two-state Markov chain with two alternatives in each state. Let the reward matrix be

$$\underline{R} = \begin{bmatrix} 10 & 4 \\ 4 & 6 \\ 14 & 3 \\ 8 & 16 \end{bmatrix} \quad \begin{cases} k = 1 \\ k = 2 \end{cases} \quad \begin{cases} i = 1 \\ i = 2 \end{cases} \quad (3.5.1)$$

Assume that the prior distribution of  $\tilde{P}$  is a matrix beta distribution



with parameter

$$\underline{\underline{m}} = \begin{bmatrix} 0.0741 & 0.0370 \\ 1.0070 & 0.3360 \\ 0.5536 & 3.9099 \\ 0.1888 & 0.1132 \end{bmatrix} \quad (3.5.2)$$

and mean

$$\underline{\underline{P}} = \begin{bmatrix} 0.687 & 0.333 \\ 0.750 & 0.250 \\ 0.125 & 0.375 \\ 0.625 & 0.375 \end{bmatrix} \quad (3.5.3)$$

Letting

$$\underline{\underline{p}} = (\tilde{p}_{11}, \tilde{p}_{12}, \tilde{p}_{11}^2, \tilde{p}_{12}^2, \tilde{p}_{21}^1, \tilde{p}_{22}^1, \tilde{p}_{21}^2, \tilde{p}_{22}^2)$$

and

$$\tilde{\underline{\underline{p}}} = E[\underline{\underline{p}}],$$

the variance-covariance matrix of this distribution is

$$E[(\underline{\underline{p}} - \tilde{\underline{\underline{p}}})(\underline{\underline{p}} - \tilde{\underline{\underline{p}}})^T] = \quad (3.5.4)$$

$$\begin{bmatrix} 0.200 & -0.200 & & & & & \\ -0.200 & 0.200 & & & & & \\ & & 0.000 & -0.000 & & & \\ & & -0.000 & 0.000 & & & \\ & & & & 0.020 & -0.020 & & \\ & & & & -0.020 & 0.020 & & \\ & & & & & & 0.180 & -0.180 \\ & & & & & & -0.180 & 0.180 \end{bmatrix}$$

Let the discount factor be  $\beta = 0.2$ .

Table 3.5.1 lists values of

$$g(n, \underline{\underline{m}}) = \begin{bmatrix} v_1(n, \underline{\underline{m}}) \\ v_2(n, \underline{\underline{m}}) \end{bmatrix} \quad (3.5.5)$$

using the terminal functions

$$v_i(\underline{\underline{m}}) = 0.000, \quad i=1, 2, \quad (3.5.6)$$



$n$	$\underline{y}(n, \underline{m})$	$\underline{\sigma}^0(n)$	$\underline{g}(n, \underline{m})$	$\Delta(n)$
0	0.000 0.000	----	10.517 13.667	20.000
1	8.002 10.999	1 2	2.515 2.668	4.000
2	10.042 13.112	1 2	0.475 0.555	0.500
3	10.418 13.562	1 2	0.099 0.105	0.160
4	10.499 13.646	1 2	0.018 0.021	0.032
5	10.514 13.664	1 2	0.003 0.003	0.006
6	10.517 13.667	1 2	0.000 0.000	0.001

$\beta = 0.2$

$\underline{y}(\underline{m}) =$  0.000  
0.000

Computation Time: 5 minutes.

Table 3.5.1



$n$	$\underline{g}(n, \underline{m})$	$\underline{\sigma}^e(n)$	$\underline{g}(n, \underline{m})$	$\Delta(n)$
0	3.750 3.750	— —	6.767 9.917	16.250
1	8.752 11.749	1 2	1.765 1.918	3.250
2	10.192 13.262	1 2	0.325 0.405	0.650
3	10.443 13.592	1 2	0.069 0.075	0.130
4	10.505 13.652	1 2	0.012 0.015	0.026
5	10.515 13.665	1 2	0.002 0.002	0.005
6	10.517 13.667	1 2	0.000 0.000	0.001

$\beta = 0.2$

$\underline{y}(\underline{m}) =$     3.750  
                              3.750

Computation time: 5 minutes.

Table 1.5.2



n	$\underline{y}(n, \underline{m})$	$\underline{\sigma}^*(n)$	$\underline{\rho}(n, \underline{m})$	$\Delta(n)$
0	10.253 13.278	---	0.265 0.390	9.747
1	10.254 13.276	1 2	0.264 0.392	1.949
2	10.470 13.586	1 2	0.048 0.082	0.390
3	10.507 13.652	1 2	0.011 0.016	0.078
4	10.516 13.665	1 2	0.002 0.003	0.016
5	10.517 13.667	1 2	0.001 0.001	0.003
6	10.518 13.669	1 2	0.000 0.000	0.001

$B = 0.2$

$\underline{y}(\underline{m}) \approx$  10.253  
13.278

Computation time: 5 minutes.



Since  $r = 3 > 0$ , the convergence is monotone increasing. It is seen that convergence to two decimal places has occurred by the sixth iteration.

The optimal initial policy vector,

$$\underline{\sigma}^*(n) = \begin{bmatrix} \sigma_1^*(n) \\ \sigma_2^*(n) \end{bmatrix} \quad (3.5.7)$$

is recorded in the third column, where  $\sigma_i^*(n)$  is the initial decision,  $k_n$ , when the system starts from state  $i$  and the  $n$ -step optimal sampling strategy is  $d(n) = (k_n, \sigma(n))$ .

Using  $x(6, \underline{M})$  as the limiting vector  $x(\underline{M})$ , the error vector of the  $n$ th iterate,  $g(n, \underline{M})$ , was computed as

$$g(n, \underline{M}) = \begin{bmatrix} 10.517 \\ 13.667 \end{bmatrix} - v(n, \underline{M}) \quad (3.5.8)$$

and is displayed in the fourth column of Table 3.5.1. The last column of the table contains the error bound

$$\Delta(n) = \beta^n \left[ \max \left\{ \frac{r}{1-\beta} - v^*, v^* - \frac{r}{1-\beta} \right\} \right]$$

defined by equation (3.3.15). Note that the bound,  $\Delta(n)$ , accurately predicts--in this example--the number of iterations required for two-place accuracy.

The computations shown in Table 3.5.2 are similar to those of Table 3.5.1, except that the terminal functions are

$$v_1(\underline{M}) = \frac{r}{1-\beta} = 3.75. \quad 101,2 \quad (3.5.9)$$

The convergence is still monotone increasing and the error functions,  $\sigma_1(n, \underline{M})$ , are reduced to approximately  $2/3$  of the corresponding values in Table 3.5.1. Five iterations are required in this case for two-place accuracy.

In Table 3.5.3, the terminal functions are the maximum expected discounted rewards when the system is operated indefinitely under a single



policy (in this case, the policy (1,2); cf. Section 4.3),

$$V(\underline{m}) = \begin{bmatrix} 10.253 \\ 13.278 \end{bmatrix} .$$

Convergence is monotonic after the first iteration. The error vector is significantly reduced as compared with corresponding entries in Tables 3.5.1 and 3.5.2. Four iterations are necessary to obtain two-place accuracy in this instance.

### 3.6 Undiscounted Processes.

When the discount factor is unity the criterion of maximizing the total expected reward over an infinite period is no longer useful since, with the possible exception of a set of sample histories of measure zero, the expected reward over an infinite period under any strategy diverges to  $+\infty$  or  $-\infty$ . An alternative criterion is to maximize the expected rate at which the process earns rewards in the steady state, or the expected gain of the process. But this criterion is not really precise since the decision-maker can, in the adaptive control process, change alternatives in any state at any time, and it is not certain that a steady state will ever be reached. Moreover, among those strategies which do lead to a steady state and which maximize the gain, there are an arbitrarily large number which are virtually equivalent—those strategies in which each alternative is sampled a large (but finite) number of times, then a fixed policy is chosen under almost perfect information. Further remarks on this class of strategies will be made in Section 5.5.

Since, for each  $\beta \in (0,1)$ , there is a well-defined criterion leading to an optimal policy, an alternative approach to the undiscounted process is to let  $\beta \rightarrow 1$  in the discounted adaptive control problem as formulated



in equation (3.1.5). For fixed  $\beta$ , let  $\underline{\Sigma}(\beta) = (\sigma_1^-(\beta), \dots, \sigma_{\bar{J}}^-(\beta))$  be an optimal initial policy, where  $\sigma_j^-(\beta)$  is the maximising value of  $k$  in equation (3.1.5) for a fixed  $\gamma \in \Gamma$ . We shall call  $\underline{\Sigma}(\beta)$  a frontier policy. If, for some  $\delta \in (0, 1)$ ,

$$\underline{\Sigma}(\beta) = \underline{\Sigma}^\delta, \quad 1 - \delta < \beta < 1 \quad (3.6.1)$$

we shall call  $\underline{\Sigma}^\delta$  an optimal initial policy for the undiscounted adaptive control problem. The existence and nature of optimal policies as defined by (3.6.1) are matters for future investigation. Blackwell [11] and Derman[16] have used this approach to undiscounted decision problems in a Markov chain with alternatives when the transition probabilities are known with certainty.



CHAPTER 4  
EXPECTED STEADY-STATE PROBABILITIES  
AND RELATED QUANTITIES

Consider a Markov chain with alternatives which is operated under a fixed policy,  $\underline{\Sigma}$ . Let  $\tilde{P}(\underline{\Sigma})$  denote the  $N \times N$  matrix of transition probabilities, assumed to have the prior distribution  $F_{\underline{\Sigma}}(P|\underline{\tau})$ .

In this chapter we examine some functions of  $\tilde{P}$  which are of importance in decision problems, with particular attention devoted to the problem of computing the means, variances, and covariances of these quantities.

Section 4.1 deals with the  $n$ -step transition probabilities and with the expected discounted reward over  $n$  transition. The second section is concerned with the steady-state probability vector. In Section 4.3 we consider the expected discounted reward over an infinite number of transitions when a fixed policy,  $\underline{\Sigma}$ , is used and, in the final section, some results concerning the expected reward per transition, or process gain, are presented. These quantities are, of course, important on their own merits; the results derived here will be applied to various terminal control models in the following chapter.

Throughout this chapter it will be assumed that a specific policy,  $\underline{\Sigma}$ , is in force and that the Markov chain is governed by the  $N \times N$  stochastic matrix,  $\tilde{P} = \tilde{P}(\underline{\Sigma})$  and that the matrix of transition rewards is  $\underline{R}(\underline{\Sigma}) = R = [r_{ij}]$ . In most cases, the dependence of various functions on  $\underline{\Sigma}$  will not be made explicit in order to simplify the notation to some extent.



#### 4.1 The n-Step Transition Probability Matrix.

If  $\underline{P}$  is a stochastic matrix governing the transitions in a Markov chain, then the probability that the system is in state  $j$  after  $n$  transitions, given that the system started in state  $i$ , is the  $(i,j)$ th element of the  $n$ th power of  $\underline{P}$ , and is denoted  $p_{ij}^{(n)}$ . When  $\tilde{P}$  is a random matrix  $\tilde{p}_{ij}^{(n)}$  is a random variable. In this section we derive expressions for the expected value of  $\tilde{p}_{ij}^{(n)}$  and the covariance of  $\tilde{p}_{ij}^{(n)}$  and  $\tilde{p}_{ij}^{(n)}$ , and examine a related quantity, the expected discounted reward over  $n$  transitions. Silver<sup>\*</sup>, using different methods, has considered the expected value of  $\tilde{p}_{ij}^{(n)}$ , assuming a matrix beta prior distribution for  $\tilde{P}$ , and has presented numerical results for a two-state process.

Theorem 4.1.1. If the prior distribution function of  $\tilde{P}$  is  $F(\underline{P}|\gamma) \in \mathcal{F}$ , a family of distributions closed under consecutive sampling, and if

$$\tilde{p}_{ij}^{(n)}(\gamma) = \int_{\mathcal{P}_N} p_{ij}^{(n)} dF(\underline{P}|\gamma) \quad i, j = 1, \dots, N \quad (4.1.1)$$

$\gamma \in \mathcal{F}$

$n=1, 2, \dots$

is the expected value of  $\tilde{p}_{ij}^{(n)}$ , then  $\tilde{p}_{ij}^{(n)}(\gamma)$  can be computed recursively from the following equations:

$$\tilde{p}_{ij}^{(n+1)}(\gamma) = \sum_{k=1}^N \tilde{p}_{ik}^{(n)}(\gamma) \tilde{p}_{kj}(\gamma), \quad i, j = 1, \dots, N \quad (4.1.2a)$$

$\gamma \in \mathcal{F}$

$n=1, 2, \dots$

$$\tilde{p}_{ij}^{(1)}(\gamma) = \tilde{p}_{ij}(\gamma) \quad i, j = 1, \dots, N \quad (4.1.2b)$$

$\gamma \in \mathcal{F}$

Proof. Since, for  $n=1, 2, \dots$ ,

$$p_{ij}^{(n+1)} = \sum_{k=1}^N p_{ik}^{(n)} p_{kj}, \quad \underline{P} \in \mathcal{P}_N$$

Lemma 2.3.2 yields

\* [38], pp. 82-87.



$$\begin{aligned}
 \tilde{p}_{ij}^{(n+1)}(\psi) &= \sum_{k=1}^N \int_{\mathcal{S}_N} p_{ik}^{(n)} p_{kj} dF(P| \psi) \\
 &= \sum_{k=1}^N \tilde{p}_{kj}(\psi) \int_{\mathcal{S}_N} p_{ik}^{(n)} dF(P| T_{kj}(\psi)) \\
 &= \sum_{k=1}^N \tilde{p}_{ik}^{(n)}(T_{kj}(\psi)) \tilde{p}_{kj}(\psi). \tag{4.1.3}
 \end{aligned}$$

Since  $p_{ij}^{(n)}$  is a continuous function of  $P$  for  $i, j = 1, \dots, N$  and  $n = 1, 2, \dots$ , the integrals in (4.1.1) and (4.1.3) exist. Q.E.D.

**Theorem 4.1.2** If the prior distribution functions of  $\tilde{P}$  is  $F(P|\psi) \in \mathcal{F}$ , a family of distributions continuous in  $\psi$ , then for  $i, j = 1, \dots, N$  and  $n = 1, 2, \dots$ ,  $\tilde{p}_{ij}^{(n)}(\psi)$  is a continuous function of  $\psi$ .

**Proof.** The theorem follows directly from Theorem 2.4.3. Q.E.D.

**Theorem 4.1.3** If  $\tilde{P}$  has the distribution function  $F(P|\psi) \in \mathcal{F}$ , a family of distributions closed under consecutive sampling, and if

$$E[\tilde{p}_{\alpha\beta}^{(n)} \tilde{p}_{\gamma\delta}^{(v)} | \psi] = \int_{\mathcal{S}_N} p_{\alpha\beta}^{(n)} p_{\gamma\delta}^{(v)} dF(P| \psi), \tag{4.1.4}$$

$\alpha, \beta, \gamma, \delta = 1, \dots, N$   
 $n, v = 1, 2, \dots$   
 $\psi \in \Xi$

then, for  $n > 1$ ,  $v > 1$ ,

$$\begin{aligned}
 E[\tilde{p}_{\alpha\beta}^{(n)} \tilde{p}_{\gamma\delta}^{(v)} | \psi] &= \sum_{k=1}^N \sum_{m=1}^N \tilde{p}_{k\beta}(\psi) \tilde{p}_{m\delta}(T_{k\beta}(\psi)) \\
 &\times E[\tilde{p}_{ak}^{(n-1)} \tilde{p}_{\gamma m}^{(v-1)} | T_{m\delta}(T_{k\beta}(\psi))], \tag{4.1.5}
 \end{aligned}$$

while, for  $n=1$  or  $v=1$ ,

$$E[\tilde{p}_{\alpha\beta} \tilde{p}_{\gamma\delta}^{(v)} | \psi] = \tilde{p}_{\alpha\beta}(\psi) \tilde{p}_{\gamma\delta}^{(v)}(T_{\alpha\beta}(\psi)) \tag{4.1.6}$$



$$E[\tilde{p}_{\alpha\beta}^{(n)} \tilde{p}_{\gamma\delta}^{(v)} | \psi] = \tilde{p}_{\alpha\beta}^{(n)}(T_{\gamma\delta}(\psi)) \tilde{p}_{\gamma\delta}^{(v)}. \quad (4.1.7)$$

**PROOF.** The functions  $\tilde{p}_{\alpha\beta}^{(n)} \tilde{p}_{\gamma\delta}^{(v)}$  are continuous on  $\mathcal{S}_N$ , hence, the integrals (4.1.4) exist. Applying Lemma 2.3.2 twice, we obtain, for the case  $n > 1, v > 1$ ,

$$\begin{aligned} E[\tilde{p}_{\alpha\beta}^{(n)} \tilde{p}_{\gamma\delta}^{(v)} | \psi] &= \sum_{k=1}^N \sum_{m=1}^N \int_{\mathcal{S}_N} p_{\alpha k}^{(n-1)} p_{k\beta}^{(v-1)} p_{\gamma m}^{(v-1)} p_{m\delta}^{(n-1)} dF(\underline{P} | \psi) \\ &= \sum_{k=1}^N \sum_{m=1}^N \bar{p}_{k\beta}(\psi) \bar{p}_{m\delta}(T_{k\beta}(\psi)) E[\tilde{p}_{\alpha k}^{(n-1)} \tilde{p}_{\gamma m}^{(v-1)} | T_{m\delta}(T_{k\beta}(\psi))]. \end{aligned}$$

If  $n = 1$ ,

$$\begin{aligned} E[\tilde{p}_{\alpha\beta} \tilde{p}_{\gamma\delta}^{(v)} | \psi] &= \int_{\mathcal{S}_N} p_{\alpha\beta} \tilde{p}_{\gamma\delta}^{(v)} dF(\underline{P} | \psi) \\ &= \bar{p}_{\alpha\beta}(\psi) \tilde{p}_{\gamma\delta}^{(v)}(T_{\alpha\beta}(\psi)) \end{aligned}$$

and similarly for  $v = 1$ . Q.E.D.

Let us now consider  $\tilde{q}_1^{(n)}(\beta, \psi)$ , the prior expected discounted reward in  $n$  transitions when the system starts in state 1 and  $F(\underline{P} | \psi)$  is the marginal prior distribution function of  $\underline{P}$ . This expectation will be required for one of the terminal control models of Chapter 5. Let  $q_1^{(n)}(\beta, \underline{P})$  be the corresponding expected discounted reward given that  $\tilde{\underline{P}} = \underline{P}$ . Then

$$q_1^{(n)}(\beta, \psi) = \int_{\mathcal{S}_N} q_1^{(n)}(\beta, \underline{P}) dF(\underline{P} | \psi). \quad (4.1.8)$$

**Theorem 4.1.4** If the prior distribution function of  $\underline{P}$  is  $F(\underline{P} | \psi) \in \mathcal{F}$ , a family of distributions closed under consecutive sampling, then  $\tilde{q}_1^{(n)}(\beta, \psi)$  can be computed recursively from the following equations:

$$\tilde{q}_1^{(n+1)}(\beta, \psi) = \sum_{k=1}^N \bar{p}_{1k}(\psi) [r_{1k} + \beta \tilde{q}_1^{(n)}(\beta, T_{1k}(\psi))] \quad (4.1.9a)$$

$i=1, \dots, N$   
 $n=1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta \leq 1$



$$\tilde{q}_j^{(1)}(\beta, \gamma) = \sum_{k=1}^N \tilde{p}_{jk}(\gamma) r_{jk} \quad \begin{array}{l} i=1, \dots, N \\ j \in F \\ 0 \leq \beta \leq 1 \end{array} \quad (4.1.9b)$$

where  $R = [r_{ij}]$  is the reward matrix.

**Proof.** For  $n=1, 2, \dots$  and all  $P \in \mathcal{P}_N$ ,  $\tilde{q}_j^{(n)}(\beta, P)$  satisfies the following renewal equation,

$$\tilde{q}_j^{(n+1)}(\beta, P) = \sum_{k=1}^N p_{jk} [r_{jk} + \beta \tilde{q}_k^{(n)}(\beta, P)]. \quad i=1, \dots, N$$

Then, using Lemma 2.3.2,

$$\begin{aligned} \tilde{q}_j^{(n+1)}(\beta, \gamma) &= \sum_{k=1}^N \tilde{p}_{jk}(\gamma) \int_{\mathcal{P}_N} [r_{jk} + \beta \tilde{q}_k^{(n)}(\beta, P)] dP | T_{jk}(\gamma) \\ &= \sum_{k=1}^N \tilde{p}_{jk}(\gamma) [r_{jk} + \beta \tilde{q}_k^{(n)}(\beta, T_{jk}(\gamma))], \end{aligned} \quad (4.1.10)$$

which is (4.1.9a). Since  $\tilde{q}_j^{(1)}(\beta, \gamma)$  is  $\tilde{q}_j(\gamma)$  as defined by equation (3.1.4), equation (4.1.9b) follows. Q.E.D.

For the case  $\beta = 1$  another method is available for evaluating  $\tilde{q}_j^{(n)}(1, \gamma)$ . In a sample of  $n$  observations let  $f_{ij}$  be the number of transitions observed from state  $i$  to state  $j$  and let  $\tilde{F} = [f_{ij}]$ , an  $N \times N$  matrix, be the transition count of the sample. Prior to the observation of the sample  $\tilde{F}$  is a random matrix and, given the initial state  $i$ , the number of transitions  $n$ , and the prior distribution of  $\tilde{P}$ , we can find the distribution of  $\tilde{F}$ , unconditional with regard to  $\tilde{P}$ . Let  $\tilde{F} = [\tilde{F}_{ij}]$  be the mean of this unconditional sampling distribution. Then

$$\tilde{q}_j^{(n)}(1, \gamma) = \sum_{j=1}^N \sum_{k=1}^N \tilde{F}_{jk} r_{jk}. \quad (4.1.11)$$

If the prior distribution of  $\tilde{P}$  is the matrix beta distribution, then the distribution of  $\tilde{F}$ , unconditional with regard to  $\tilde{P}$ , is the beta-Whittle



distribution, which is discussed in Section 6.5.

#### 4.2 The Steady-State Probability Vector.

Let  $\underline{P}$  be an ergodic stochastic matrix. Then there is associated with  $\underline{P}$  a unique vector of steady-state probabilities,  $\underline{\pi}(\underline{P}) = (\pi_1, \dots, \pi_N)$ , where  $\pi_i = \pi_i(\underline{P})$  is the steady-state probability that the system is in state  $i$  ( $i = 1, \dots, N$ ). The vector  $\underline{\pi}$  satisfies the following system of simultaneous equations,

$$\underline{\pi} = \underline{\pi}\underline{P} \quad (4.2.1a)$$

$$\sum_{i=1}^N \pi_i = 1. \quad (4.2.1b)$$

If  $\tilde{\underline{P}}$  is a random matrix with an arbitrary distribution function,  $F(\underline{P} | \mathcal{U})$ , which satisfies a mild continuity condition, we show below that the subset of non-ergodic matrices in  $\mathcal{S}_N$  is a set of measure zero. Thus, it is meaningful to speak of the random vector  $\tilde{\underline{\pi}}$ .

We are chiefly concerned, in this section, with the expected value,  $\underline{\pi}(\mathcal{U}) = (\bar{\pi}_1(\mathcal{U}), \dots, \bar{\pi}_N(\mathcal{U}))$ , of  $\tilde{\underline{\pi}}$ . It is shown that this expectation exists and that  $\lim_{n \rightarrow \infty} \underline{p}_{ij}^{(n)}(\mathcal{U}) = \bar{\pi}_j(\mathcal{U})$ . We then assume that  $F(\underline{P} | \mathcal{U}) \in \mathcal{F}$ , a family of distributions closed under the consecutive sampling rule, and derive a functional equation for  $\underline{\pi}(\mathcal{U})$ . Methods of successive approximations based on this equation are discussed, then some numerical results are presented. We conclude with a discussion of the covariance of  $\tilde{\pi}_i$  and  $\tilde{\pi}_j$ .

##### 4.2.1 Existence of the Moments of $\tilde{\underline{\pi}}$ .

Let us now consider conditions on the prior distribution of  $\tilde{\underline{P}}$  which insure the existence of the general joint moment of the elements of  $\tilde{\underline{\pi}}$ ,



$$E \left[ \prod_{i=1}^N \tilde{\pi}_i^{v_i} | \psi \right] = \int_{\mathcal{S}_N} \prod_{i=1}^N \pi_i^{v_i} dF(P|\psi), \quad \psi \in \Psi \quad (4.2.2)$$

where the  $v_i$  are nonnegative integers.

Let

$$\mathcal{S}_N^* = \left\{ P \mid P \in \mathcal{S}_N, \quad 0 < p_{ij} < 1 \quad (i, j = 1, \dots, N) \right\} \quad (4.2.3)$$

be the set of all positive stochastic matrices and, for  $0 < a < 1$ , define the set

$$\mathcal{S}_N^a = \left\{ P \mid P \in \mathcal{S}_N^*, \quad a \leq p_{ij} \leq 1-a \quad (i, j = 1, \dots, N) \right\}. \quad 0 < a < 1 \quad (4.2.4)$$

We remark that  $\mathcal{S}_N^a$  is a closed and bounded, hence, compact, subset of  $E_N^2$  and that  $\mathcal{S}_N^a \subset \mathcal{S}_N^* \subset \mathcal{S}_N$  for any  $a$  in the open interval  $(0, 1)$ .

For fixed  $a \in (0, 1)$ , let  $S(a)$  be any subset of  $E_N^2$  such that

$$\mathcal{S}_N - \mathcal{S}_N^a \neq S(a) \quad (4.2.5a)$$

and

$$\mathcal{S}_N - \mathcal{S}_N^a \subset S(a). \quad (4.2.5b)$$

Thus, for all  $a \in (0, 1)$ , the boundary of  $\mathcal{S}_N^a$  is a proper subset of  $S(a)$ .

If, for some  $a \in (0, 1)$ , there exists a set  $S(a)$  satisfying (4.2.5) such that  $F(P|\psi)$ , the prior distribution function of  $\underline{P}$ , is continuous on  $S(a)$ , then  $F(P|\psi)$  is said to be continuous on the boundary of  $\mathcal{S}_N^a$ . If  $\mathcal{F}$  is a family of distributions indexed by  $\psi$  every member of which is continuous on the boundary of  $\mathcal{S}_N^a$ , then  $\mathcal{F}$  is also said to be continuous on the boundary of  $\mathcal{S}_N^a$ .

The following lemma shows that continuity of  $F(P|\psi)$  on the boundary of  $\mathcal{S}_N^a$  is a necessary and sufficient condition for the set of boundary points,  $\mathcal{S}_N - \mathcal{S}_N^a$ , to be a set of measure zero. If  $P \in \mathcal{S}_N^*$ , then  $P$  consists of a single chain with no transient states. Thus, the subset of



$\mathcal{S}_N$  which includes all periodic and multiple-chain transition matrices, as well as those single-chain transition matrices which have transient states, is contained within  $\mathcal{S}_N - \mathcal{S}_N^*$ . The import of Lemma 4.2.1 is that, provided  $F(P|\gamma)$  is continuous on the boundary of  $\mathcal{S}_N$ , we need only consider transition matrices in  $\mathcal{S}_N^*$ . In this case, with probability one,  $\tilde{\pi}$  exists and, moreover,  $\tilde{\pi}_j > 0$  ( $j=1, \dots, N$ ).

**Lemma 4.2.1** If  $F(P|\gamma)$  is the prior distribution function of  $\tilde{P}$ , then a necessary and sufficient condition that  $\mathcal{S}_N - \mathcal{S}_N^*$  be a set of measure zero relative to the prior distribution is that  $F(P|\gamma)$  be continuous on the boundary of  $\mathcal{S}_N$ .

Proof. For all  $a \in (0,1)$ , define the sets

$$c_{ij}(a) = \left\{ P \mid P \in \mathcal{S}_N^*, 0 \leq p_{ij} \leq a \right\}. \quad \begin{matrix} i,j=1, \dots, N \\ 0 < a < 1 \end{matrix} \quad (4.2.6)$$

Then

$$\mathcal{S}_N - \mathcal{S}_N^* \subset \mathcal{S}_N - \mathcal{S}_N^a \subset \bigcup_{i,j=1}^N c_{ij}(a), \quad 0 < a < 1 \quad (4.2.7)$$

and, for all  $a \in (0,1)$ , the probability measure of the set  $\mathcal{S}_N - \mathcal{S}_N^a$  has the bound

$$\int_{\mathcal{S}_N - \mathcal{S}_N^a} dF(P|\gamma) \leq \sum_{i=1}^N \sum_{j=1}^N \int_{c_{ij}(a)} dF(P|\gamma). \quad \begin{matrix} i,j=1, \dots, N \\ 0 < a < 1 \end{matrix} \quad (4.2.8)$$

If  $F_{ij}(a|\gamma)$  is the marginal distribution function of  $\tilde{P}_{ij}$ , then

$$\int_{c_{ij}(a)} dF(P|\gamma) = F_{ij}(a|\gamma). \quad \begin{matrix} i,j=1, \dots, N \\ 0 < a < 1 \end{matrix} \quad (4.2.9)$$

Assume that, for fixed  $a \in (0,1)$ , there exists a set  $S(a)$  satisfying (4.2.5) on which  $F(P|\gamma)$  is continuous. Let  $\epsilon > 0$  be given. Since  $F(P|\gamma)$  is continuous on  $S(a)$  we may choose on  $a'$  such that  $0 < a' < a$  and



$$F_{ij}(a^*|\gamma) < \frac{\epsilon}{N^2}. \quad i,j=1, \dots, N \quad (4.2.10)$$

Then

$$\int_{\mathcal{S}_N - \mathcal{S}_N^*} dF(P|\gamma) \leq \sum_{i=1}^N \sum_{j=1}^N F_{ij}(a^*|\gamma) < \epsilon \quad (4.2.11)$$

and, since  $\epsilon$  is arbitrary,  $\mathcal{S}_N - \mathcal{S}_N^*$  is a set of measure zero, proving sufficiency.

To demonstrate necessity it suffices to note that, if there does not exist a set  $S(a)$  which satisfies the conditions of the lemma, then  $F(P|\gamma)$  must assign positive probability to at least one of the boundary points of  $\mathcal{S}_N$ . Q.E.D.

We remark that, in the case of the matrix beta distribution, the existence of a density function implies that the corresponding distribution function is continuous on  $E_N$ . Therefore, the family of matrix beta distributions is continuous on the boundary of  $\mathcal{S}_N$ .

**Theorem 4.2.2** If the prior distribution function of  $\tilde{P}$ ,  $F(\tilde{P}|\gamma)$ , is continuous on the boundary of  $\mathcal{S}_N$ , then the joint moments (4.2.2) exist for all nonnegative integers  $v_i$ . If, furthermore,  $F(\tilde{P}|\gamma) \in \mathcal{F}$ , a family of distributions continuous in  $\gamma$ , then  $E[\prod_{i=1}^N \tilde{\pi}_i^{v_i} | \gamma]$  is a continuous function of  $\gamma$ .

**Proof.** By Lemma 4.2.1,

$$E[\prod_{i=1}^N \tilde{\pi}_i^{v_i} | \gamma] = \int_{\mathcal{S}_N} \prod_{i=1}^N \tilde{\pi}_i^{v_i} dF(\tilde{P}|\gamma). \quad (4.2.12)$$

Let  $D_{jj}(P)$  be the cofactor of the  $j$ th diagonal elements of the matrix

$$D(P) = [P_n^T - I]. \quad (4.2.13)$$



It can be shown<sup>\*</sup> that, for all  $P \in \mathcal{S}_N^*$ ,

$$\pi_j(P) = \frac{D_{jj}(P)}{\sum_{k=1}^N D_{kk}(P)}. \quad j=1, \dots, N \quad (4.2.14)$$

Since  $D_{jj}(P)$  is a sum of products involving elements of  $P$ ,  $\pi_j(P)$  is a continuous bounded function of  $P$  on  $\mathcal{S}_N^*$  and the integral (4.2.12) exists. When  $F(P|\gamma) \in \mathcal{F}_1$ , a family of distributions continuous in  $\gamma$ , continuity of  $E\left[\prod_{i=1}^N \pi_i^{u_i} | \gamma\right]$  follows from Theorem 2.4.3. Q.E.D.

Theorem 4.2.2 can be proved under the weaker condition that the set of nonergodic transition probability matrices in  $\mathcal{S}_N$  is a set of measure zero, but this criterion is more difficult to apply in practice than that of continuity on the boundary. There are many problems, however, in which it is necessary to assign positive probability to ergodic matrices on the boundary of  $\mathcal{S}_N$ . For example, in some random walk models  $\tilde{P}$  is known to be a Jacobi matrix—that is,  $\tilde{p}_{ij} = 0$  with probability one if  $|i - j| > 1$ . In this case the theory presented here can be applied by assigning a prior distribution to the  $N \times 3$  generalized stochastic matrix  $\tilde{P}$ , where the  $i$ th row of  $\tilde{P}$  consists of the elements  $\tilde{p}_{i,i-1}, \tilde{p}_{ii}, \tilde{p}_{i,i+1}$ . This technique can be applied to any ergodic transition probability matrix in which some elements are known to be zero.

**4.2.2 An Ergodic Theorem.** We now establish that, if  $\tilde{p}_{ij}^{(n)}(\gamma)$  is the mean  $n$ -step transition probability defined by equation (4.1.1),

---

\* Singer [39]. This result was apparently first discovered by Kihoc in 1934 in his doctoral thesis (Rumanian). Cf. Rosenblatt [34].



$$\lim_{n \rightarrow \infty} p_{ij}^{(n)}(\gamma) = \pi_j(\gamma), \quad \begin{matrix} i,j=1, \dots, N \\ \gamma \in \mathbb{S}_N^{\alpha} \end{matrix} \quad (4.2.15)$$

where  $\underline{\pi}(\gamma) = (\pi_1(\gamma), \dots, \pi_N(\gamma))$  is the expected value of  $\underline{P}$ . The only assumption which is made about  $P(\underline{P}|\gamma)$ , the prior distribution function of  $\underline{P}$ , is that  $P(\underline{P}|\gamma)$  is continuous on the boundary of  $\mathcal{S}_N^{\alpha}$ . In order to establish (4.2.15) it must first be shown that, for any fixed  $a \in (0,1)$ ,  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j(P)$  uniformly in  $P$  on  $\mathcal{S}_N^{\alpha}$ . This is the content of the following two lemmas.

**Lemma 4.2.3** For some fixed  $a \in (0,1)$  let

$$\Delta(P) = \min_{\underline{P} \in \mathcal{S}_N^{\alpha}} \{P_{ij}\}, \quad P \in \mathcal{S}_N^{\alpha} \quad (4.2.16)$$

be a function of  $P$  on  $\mathcal{S}_N^{\alpha}$ . Then  $\Delta(P)$  is continuous on  $\mathcal{S}_N^{\alpha}$ .

**Proof.** Let  $\epsilon > 0$  be given. It must be shown that, for any fixed  $P \in \mathcal{S}_N^{\alpha}$ , there exists a  $\delta > 0$  such that  $|\Delta(P) - \Delta(Q)| < \epsilon$  whenever  $Q \in \mathcal{S}_N^{\alpha}$  and  $\|P - Q\| < \delta$ . Let

$$p_{kk} = \min_{1,j} \{p_{kj}\}$$

and let

$$\epsilon' = p_{kk} - p_{ij} > 0 \quad (4.2.17)$$

where  $p_{kk}$  is the smallest element of  $P$  not equal to  $p_{ij}$  (assuming, for the moment, that such an element exists). Choose  $\delta = \min[\epsilon, \epsilon'/2]$ .

Then, for any  $Q \in \mathcal{S}_N^{\alpha}$ , if  $\|P - Q\| < \delta$ , we have

$$|p_{ij} - p_{ij}| < \delta \leq \epsilon'/2 \quad i,j=1, \dots, N$$

and

\* This result can be proved under the weaker condition that the set of nonergodic matrices in  $\mathcal{S}_N$  is a set of measure zero, using the bounded convergence theorem of measure theory. The proof given here brings out some interesting features of the convergence of  $p_{ij}^{(n)}$  to  $\pi_j(P)$  on  $\mathcal{S}_N$  and does not require a knowledge of measure theory.



$$p_{ij} - \epsilon^*/2 < q_{ij} < p_{ij} + \epsilon^*/2. \quad i,j=1, \dots, n \quad (4.2.18)$$

Let

$$S_{\underline{ij}} = \{(i,j) \mid p_{ij} \neq p_{\underline{ij}}, \quad i,j=1, \dots, n\}, \quad (4.2.19)$$

assumed, for the moment, to be nonempty. Then, using (4.2.17) and the definition of  $p_{\underline{ij}}$ ,

$$p_{\underline{ij}} = p_{\underline{nn}} - \epsilon^* \leq p_{ij} = \epsilon^*. \quad (i,j) \in S_{\underline{ij}} \quad (4.2.20)$$

Thus, by (4.2.18),

$$q_{ij} < p_{ij} + \epsilon^*/2 \leq p_{ij} - \epsilon^*/2 < q_{\underline{ij}}. \quad (4.2.21)$$

$(i,j) \in S_{\underline{ij}}$

If  $q_{\beta\gamma} = \min_{i,j} \{q_{ij}\}$ , (4.2.21) implies that  $(\beta,\gamma) \in S_{\underline{ij}}$ , and, therefore, that  $p_{\underline{ij}} = p_{\beta\gamma}$ . Thus,

$$|\Delta(\underline{P}) - \Delta(\underline{Q})| = |p_{\beta\gamma} - q_{\beta\gamma}| < \delta \leq \epsilon. \quad (4.2.22)$$

Suppose now that there is no smallest element of  $\underline{P}$  not equal to  $p_{\underline{ij}}$ . Then  $S_{\underline{ij}}$  is empty and  $p_{ij} = \frac{1}{n}$  ( $i,j=1, \dots, n$ ). Choosing  $\delta = \epsilon$ ,  $\|\underline{P} - \underline{Q}\| < \delta$  implies that

$$\left| \frac{1}{n} - q_{ij} \right| < \epsilon, \quad i,j=1, \dots, n \quad (4.2.23)$$

and, hence, that

$$|\Delta(\underline{P}) - \Delta(\underline{Q})| < \epsilon, \quad (4.2.24)$$

proving the lemma. Q.E.D.

**Lemma 4.2.4** Let  $\alpha \in (0,1)$  be fixed. Then, for  $i,j = 1, \dots, n$ ,

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j(P) \quad (4.2.25)$$

uniformly in  $P$  on  $\mathcal{A}_N^\alpha$ .

**Proof.** Let  $\epsilon > 0$  be given. It must be shown that there exists a



positive integer  $v_a$  such that, if  $n > v_a$ ,

$$\left| p_{ij}^{(n)} - \pi_j(p) \right| < \epsilon \quad (4.2.26)$$

for all  $p \in \mathcal{S}_N^a$ . This will be done by showing that the sequence of functions  $\{p_{ij}^{(n)} - \pi_j(p)\}$  is bounded by a sequence of functions of  $p$  which goes to zero uniformly on  $\mathcal{S}_N^a$  as  $n \rightarrow \infty$ . Define the functions

$$\Delta(p) = \min_{1,j} \{p_{ij}\}, \quad p \in \mathcal{S}_N^a \quad (4.2.27)$$

and

$$d_n(p) = [1 - 2\Delta(p)]^n \quad (4.2.28)$$

$n=1, 2, \dots$   
 $p \in \mathcal{S}_N^a$

Since  $0 < \Delta(p) \leq \frac{1}{2}$  for any  $p \in \mathcal{S}_N^a$ ,

$$0 \leq 1 - 2\Delta(p) < 1 \quad p \in \mathcal{S}_N^a \quad (4.2.29)$$

and

$$d_n(p) \geq d_{n+1}(p). \quad n=1, 2, \dots \quad (4.2.30)$$

By Lemma 4.2.3 and equations (4.2.29) and (4.2.30),  $\{d_n(p)\}$  is a monotonically decreasing sequence of functions which are continuous on the compact set  $\mathcal{S}_N^a$ , the sequence converging to the continuous function zero. Therefore\*,  $\{d_n(p)\} \rightarrow 0$  uniformly on  $\mathcal{S}_N^a$ . It is easily established that

$$\left| p_{ij}^{(n)} - \pi_j(p) \right| \leq d_n(p). \quad n=1, 2, \dots \quad (4.2.31)$$

By choosing a positive integer  $v_a$  such that  $0 \leq d_n(p) < \epsilon$  for all  $n > v_a$  and  $p \in \mathcal{S}_N^a$ , (4.2.26) is obtained. Q.E.D.

\* Rudin [35], p. 136.

# Kemeny and Snell [26], p. 71.



**Theorem 4.2.5** Let  $F(P|t)$  be the prior distribution function of the random  $N \times N$  stochastic matrix  $\tilde{P}$  and let

$$\tilde{p}_{ij}^{(n)}(t) = \int_{\mathcal{S}_N^n} p_{ij}^{(n)} dF(P|t) \quad \begin{array}{l} i,j=1, \dots, N \\ n=1, 2, \dots \\ t \in \mathbb{R} \end{array} \quad (4.2.32)$$

$$\bar{\pi}_j(t) = \int_{\mathcal{S}_N^n} \pi_j(P) dF(P|t). \quad \begin{array}{l} j=1, \dots, N \\ t \in \mathbb{R} \end{array} \quad (4.2.33)$$

Then, if  $F(P|t)$  is continuous on the boundary of  $\mathcal{S}_N^n$ ,

$$\lim_{n \rightarrow \infty} \tilde{p}_{ij}^{(n)}(t) = \bar{\pi}_j(t). \quad \begin{array}{l} i,j=1, \dots, N \\ t \in \mathbb{R} \end{array} \quad (4.2.34)$$

**Proof.** Let  $\epsilon > 0$  be given. For any  $a \in (0, 1)$ ,

$$|\tilde{p}_{ij}^{(n)}(t) - \bar{\pi}_j(t)| \leq \int_{\mathcal{S}_N^a} |\tilde{p}_{ij}^{(n)}(P) - \pi_j(P)| dF(P|t) + \int_{\mathcal{S}_N - \mathcal{S}_N^a} |\tilde{p}_{ij}^{(n)}(P) - \pi_j(P)| dF(P|t) \quad (4.2.35)$$

Let  $C_{km}(a)$  be defined by equation (4.2.6) and let  $F_{km}(P|t)$  be the marginal distribution function of  $\tilde{p}_{km}^{(n)}$ . Then, noting that  $|\tilde{p}_{ij}^{(n)} - \pi_j(P)| \leq C_{ij}(a)$ , the second integral of (4.2.35) has the bound

$$\int_{\mathcal{S}_N - \mathcal{S}_N^a} |\tilde{p}_{ij}^{(n)} - \pi_j(P)| dF(P|t) \leq \sum_{k=1}^N \sum_{m=1}^N F_{km}(a|t). \quad (4.2.36)$$

$0 < a < 1$

Since  $F(P|t)$  is continuous on the boundary of  $\mathcal{S}_N^n$ , there is an  $a^* \in (0, 1)$  and a set  $S(a^*)$  satisfying equation (4.2.5) such that  $F(P|t)$  is continuous on  $S(a^*)$ . In equation (4.2.35) let  $a < a^*$  be chosen such that

$F_{km}(a|t) < \epsilon/2N^2$  for  $k, m = 1, \dots, N$ . Then

$$\int_{\mathcal{S}_N - \mathcal{S}_N^a} |\tilde{p}_{ij}^{(n)} - \pi_j(P)| dF(P|t) < \frac{\epsilon}{2}. \quad i=1, 2, \dots \quad (4.2.37)$$

Having fixed  $a$ , choose a positive integer  $v_a$  such that

$$|\tilde{p}_{ij}^{(n)} - \pi_j(P)| < \frac{\epsilon}{2}$$



for all  $n > v_a$  and all  $P \in \mathcal{P}_N^G$ . Then

$$\int_{\mathcal{P}_N^G} \left| p_{1j}^{(n)} - \bar{\pi}_j(P) \right| dF(P|\psi) < \frac{\epsilon}{2} \quad n > v_a \quad (4.2.38)$$

and equations (4.2.35), (4.2.37), and (4.2.38) yield

$$\left| \bar{p}_{1j}^{(n)}(\psi) - \bar{\pi}_j(\psi) \right| < \epsilon. \quad n > v_a \quad (4.2.39)$$

Q.E.D.

Theorem 4.2.5 shows that we can approximate  $\bar{\pi}_j(\psi)$  by  $\bar{p}_{1j}^{(n)}(\psi)$  using equation (4.1.2). A recursive program was written to carry out this approximation when  $\bar{P}$  has a matrix beta distribution.\* Some sample computations of  $E[\bar{P}^n | \underline{m}]$  are displayed in Tables 4.2.1-4.2.3. We have shown at the base of each table the parameter  $\underline{m}$  of the prior distribution of  $\bar{P}$  and the mean,  $\bar{P}$ , of this distribution. The matrix  $\underline{V}(\bar{P})$  which appears below the table has as its  $(i,j)$ th element the prior variance of  $\bar{p}_{ij}$ . Silver [38] has conjectured that  $\bar{\pi}_j(\psi)$  can be approximated reasonably well by  $\underline{\pi}(\bar{P})$ , the steady-state probability vector corresponding to  $\bar{P}(\underline{m})$ , the mean of the prior distribution. This approximation is also given with each table. All work was performed on an IBM 7094 computer.

In Table 4.2.1, where a  $2 \times 2$  transition matrix is considered, it is seen that  $\bar{\pi}_j(\underline{m})$  is obtained with three-place accuracy for  $n = 5$  and with five-place accuracy for  $n = 8$ . In this instance,  $\bar{p}_{1j}^{(n)}(\underline{m})$  converges monotonically to  $\bar{\pi}_j(\underline{m})$ . The total time required to compute the eight entries of Table 4.2.1 was 0.70 minutes.

In Table 4.2.2, a  $2 \times 2$  transition matrix is treated which has prior variances which are larger than those of the matrix considered in Table 4.2.1. In this instance convergence of  $\bar{p}_{1j}^{(n)}(\underline{m})$  to  $\bar{\pi}_j(\underline{m})$  is much slower

---

\* See Appendix C for the program listing.



$n$	$E[\underline{P} \underline{M}]$
1	$\begin{bmatrix} 0.93454 & 0.06546 \\ 0.86357 & 0.13643 \end{bmatrix}$
2	$\begin{bmatrix} 0.93370 & 0.06630 \\ 0.92027 & 0.07973 \end{bmatrix}$
3	$\begin{bmatrix} 0.93297 & 0.06703 \\ 0.93096 & 0.06904 \end{bmatrix}$
4	$\begin{bmatrix} 0.93296 & 0.06704 \\ 0.93244 & 0.06756 \end{bmatrix}$
5	$\begin{bmatrix} 0.93293 & 0.06707 \\ 0.93283 & 0.06717 \end{bmatrix}$
6	$\begin{bmatrix} 0.93293 & 0.06707 \\ 0.93290 & 0.06720 \end{bmatrix}$
7	$\begin{bmatrix} 0.93293 & 0.06707 \\ 0.93292 & 0.06708 \end{bmatrix}$
8	$\begin{bmatrix} 0.93293 & 0.06707 \\ 0.93293 & 0.06707 \end{bmatrix}$

$$\underline{m} = \begin{bmatrix} 24.105 & 0.988 \\ 21.313 & 3.367 \end{bmatrix}$$

$$\underline{\hat{P}} = \begin{bmatrix} 0.93454 & 0.06546 \\ 0.86357 & 0.13643 \end{bmatrix}$$

$$\underline{\pi}(\underline{P}) = (0.92954 \quad 0.07046)$$

$$\underline{v}(\underline{\hat{P}}) = \begin{bmatrix} 0.0038 & 0.0030 \\ 0.0048 & 0.0048 \end{bmatrix}$$

Computation Time: 0.70 minutes.

Table 4.2.1



$n$	$E[\tilde{P}^n   \mathcal{M}]$
1	$\begin{bmatrix} 0.41625 & 0.58375 \\ 0.35521 & 0.64479 \end{bmatrix}$
2	$\begin{bmatrix} 0.53220 & 0.46780 \\ 0.29321 & 0.70679 \end{bmatrix}$
3	$\begin{bmatrix} 0.43070 & 0.56930 \\ 0.39086 & 0.61914 \end{bmatrix}$
4	$\begin{bmatrix} 0.47832 & 0.52168 \\ 0.37443 & 0.65257 \end{bmatrix}$
5	$\begin{bmatrix} 0.43084 & 0.56916 \\ 0.39157 & 0.60849 \end{bmatrix}$
6	$\begin{bmatrix} 0.45896 & 0.54104 \\ 0.36973 & 0.63027 \end{bmatrix}$
7	$\begin{bmatrix} 0.42989 & 0.57011 \\ 0.39759 & 0.60241 \end{bmatrix}$
8	$\begin{bmatrix} 0.44918 & 0.55082 \\ 0.38179 & 0.61821 \end{bmatrix}$
9	$\begin{bmatrix} 0.42895 & 0.57105 \\ 0.40150 & 0.59850 \end{bmatrix}$
10	$\begin{bmatrix} 0.44333 & 0.55667 \\ 0.39931 & 0.61069 \end{bmatrix}$
11	$\begin{bmatrix} 0.42815 & 0.57185 \\ 0.40424 & 0.59576 \end{bmatrix}$
12	$\begin{bmatrix} 0.43945 & 0.56055 \\ 0.39445 & 0.60555 \end{bmatrix}$
13	$\begin{bmatrix} 0.42749 & 0.57251 \\ 0.40629 & 0.59371 \end{bmatrix}$

$$\mathcal{M} = \begin{bmatrix} 0.251 & 0.352 \\ 0.616 & 1.120 \end{bmatrix}$$

$$\mathbb{E}(\tilde{P}) = (0.37830 \quad 0.62170)$$

$$\tilde{P} = \begin{bmatrix} 0.41625 & 0.58375 \\ 0.35521 & 0.64479 \end{bmatrix}$$

$$V(\tilde{P}) = \begin{bmatrix} 0.1516 & 0.1516 \\ 0.0837 & 0.0837 \end{bmatrix}$$

Computation Time: 5.75 minutes.

Table 4.2.2



$n$	$E[\tilde{P}^n   \mathcal{M}]$		
2	$\begin{bmatrix} 0.60196 \\ 0.19754 \\ 0.19297 \end{bmatrix}$	$\begin{bmatrix} 0.21244 \\ 0.55168 \\ 0.32211 \end{bmatrix}$	$\begin{bmatrix} 0.18560 \\ 0.25078 \\ 0.48492 \end{bmatrix}$
4	$\begin{bmatrix} 0.45155 \\ 0.27034 \\ 0.26785 \end{bmatrix}$	$\begin{bmatrix} 0.29867 \\ 0.43724 \\ 0.36814 \end{bmatrix}$	$\begin{bmatrix} 0.24978 \\ 0.29222 \\ 0.36401 \end{bmatrix}$
6	$\begin{bmatrix} 0.38841 \\ 0.30102 \\ 0.29958 \end{bmatrix}$	$\begin{bmatrix} 0.33563 \\ 0.39832 \\ 0.37440 \end{bmatrix}$	$\begin{bmatrix} 0.27536 \\ 0.30016 \\ 0.32602 \end{bmatrix}$
8	$\begin{bmatrix} 0.35970 \\ 0.31487 \\ 0.31410 \end{bmatrix}$	$\begin{bmatrix} 0.35256 \\ 0.38373 \\ 0.37405 \end{bmatrix}$	$\begin{bmatrix} 0.28774 \\ 0.30140 \\ 0.31185 \end{bmatrix}$
9	$\begin{bmatrix} 0.35155 \\ 0.31896 \\ 0.31802 \end{bmatrix}$	$\begin{bmatrix} 0.35691 \\ 0.37371 \\ 0.38094 \end{bmatrix}$	$\begin{bmatrix} 0.29154 \\ 0.30733 \\ 0.30104 \end{bmatrix}$

$$\underline{m} = \begin{bmatrix} 18.265 \\ 2.395 \\ 1.005 \end{bmatrix} \quad \begin{bmatrix} 2.102 \\ 5.168 \\ 7.612 \end{bmatrix} \quad \begin{bmatrix} 3.910 \\ 10.111 \\ 1.212 \end{bmatrix}$$

$$\tilde{P} = \begin{bmatrix} 0.75236 \\ 0.13502 \\ 0.10225 \end{bmatrix} \quad \begin{bmatrix} 0.08650 \\ 0.29257 \\ 0.77444 \end{bmatrix} \quad \begin{bmatrix} 0.16106 \\ 0.57241 \\ 0.12331 \end{bmatrix}$$

$$\underline{\Pi}(\tilde{P}) = (0.32697 \quad 0.37084 \quad 0.30219)$$

$$\underline{V}(\tilde{P}) = \begin{bmatrix} 0.0074 \\ 0.0063 \\ 0.0085 \end{bmatrix} \quad \begin{bmatrix} 0.0031 \\ 0.0111 \\ 0.0161 \end{bmatrix} \quad \begin{bmatrix} 0.0053 \\ 0.0131 \\ 0.0100 \end{bmatrix}$$

Computation Time: 5 minutes ( $n = 1, \dots, 8$ )  
 3.26 minutes ( $n = 9$ ).

Table 4.2.3



and is not monotonic. For  $n = 13$ ,  $\tilde{p}_{11}^{(n)}(\underline{m})$  and  $p_{21}^{(n)}(\underline{m})$  agree only in the first decimal place. The 13 entries of this table took a total of 5.75 minutes to compute.

Some sample computations for a three-state process are shown in Table 4.2.3. Five minutes were required to compute the first eight entries in this case. The computation of  $E[\tilde{P}^9 | \underline{m}]$  required 3.26 minutes. Convergence is slow and is not monotonic.

We remark that the computation time of  $\tilde{p}_{1j}^{(n)}(\underline{m})$  increases exponentially with  $n$  and linearly with  $N$ .

**4.2.3 Successive Approximations.** The numerical calculation of  $\bar{\pi}_j(\gamma)$  is a problem of some difficulty. To obtain an explicit formula for  $\bar{\pi}_j(\gamma)$  in terms of  $\gamma$  is even more difficult; this general problem has not yet been solved. Silver [38], assuming a matrix beta distribution for  $\tilde{P}$ , has calculated  $\bar{\pi}_j(\underline{m})$  for various parameters,  $\underline{m}$ , using Monte Carlo techniques. He has also shown that, for a two-state chain with one row of  $\tilde{P}$  known with certainty and a beta distribution on the other row, the expected value of  $\tilde{\pi}_j$  is a Gaussian hypergeometric function. This result is generalised in Section 8.5, where a series expansion of  $\bar{\pi}_j(\underline{m})$  is obtained when the  $2 \times 2$  random matrix  $\tilde{P}$  has the matrix beta distribution with parameter  $\underline{m}$ .

One method of computing  $\bar{\pi}_j(\gamma)$  is to use the ergodic theorem of the last section. A more general basis for the calculation of  $\bar{\pi}_j(\gamma)$  is provided in the next theorem.

**Theorem 4.2.6** If  $\tilde{P}$  has the distribution function  $F(\underline{P} | \gamma) \in \mathcal{F}_1$ , where  $\mathcal{F}_1$  is closed under consecutive sampling and is continuous on the



boundary of  $\mathcal{S}_N$ , then the expectations  $\bar{\pi}_j(\psi)$  simultaneously satisfy the functional equations

$$\bar{\pi}_j(\psi) = \sum_{k=1}^N \bar{\pi}_k(\tau_{kj}(\psi)) \bar{p}_{kj}(\psi), \quad j=1, \dots, N \quad (4.2.40a)$$

together with

$$\sum_{j=1}^N \bar{\pi}_j(\psi) = 1. \quad \psi \in \Psi \quad (4.2.40b)$$

Remark. The condition (4.2.40b) is necessary to insure a unique solution to (4.2.40a) since, if  $\underline{\pi}(\psi)$  satisfies (4.2.40a),  $e\underline{\pi}(\psi)$  also satisfies (4.2.40a) for all real numbers  $e$ . Even with this additional constraint we have been unable to prove that  $E[\bar{\pi}|\psi]$  is the unique solution to (4.2.40), although we conjecture that this is true.

Proof. For  $j=1, \dots, N$  and  $\psi \in \Psi$ , using (4.2.1a) and Lemma 2.3.2,

$$\begin{aligned} \bar{\pi}_j(\psi) &= \sum_{k=1}^N \int_{\mathcal{S}_N} \pi_k(p) p_{kj} dF(p|\psi) \\ &= \sum_{k=1}^N \bar{\pi}_k(\tau_{kj}(\psi)) \bar{p}_{kj}(\psi), \end{aligned} \quad (4.2.41)$$

which is (4.2.40a). Summing (4.2.33) over  $j$  yields (4.2.40b). The integrals involved exist by virtue of Lemma 4.2.1 and the continuity of  $\pi_j(p)$  on  $\mathcal{S}_N^*$ . Q.E.D.

Let the vector function,  $\bar{\pi}(n, \psi) = (\bar{\pi}_1(n, \psi), \dots, \bar{\pi}_N(n, \psi))$ , be defined by the equations

$$\begin{aligned} \bar{\pi}_j(n+1, \psi) &= \sum_{k=1}^N \bar{\pi}_k(n, \tau_{kj}(\psi)) \bar{p}_{kj}(\psi) \\ &\quad j=1, \dots, N \\ &\quad n=1, 2, 3, \dots \\ &\quad \psi \in \Psi \end{aligned} \quad (4.2.42)$$

together with a terminal function  $\bar{\pi}(0, \psi) = (\bar{\pi}_1(0, \psi), \dots, \bar{\pi}_N(0, \psi))$



which satisfies the conditions

$$0 \leq \bar{\pi}_j(0, \gamma) \leq 1 \quad i=1, \dots, N \quad (\text{4.2.43a})$$

$$\gamma \in \Psi$$

$$\sum_{i=1}^N \bar{\pi}_i(0, \gamma) = 1. \quad \gamma \in \Psi \quad (\text{4.2.43b})$$

The function  $\tilde{p}_{kj}(\gamma)$  is the  $(k, j)$ th element of the expected value of  $\tilde{P}$  when  $F(\underline{P} | \gamma)$  is the prior distribution function of  $\tilde{P}$ .

If  $\lim_{n \rightarrow \infty} \underline{\pi}(n, \gamma)$  exists, then this limit satisfies equation (4.2.40a). In general,  $C(n, \gamma) \geq \sum_{j=1}^N \bar{\pi}_j(n, \gamma)$  does not equal unity and, therefore,  $\lim_{n \rightarrow \infty} C(n, \gamma)$  need not equal unity. However, if  $\lim_{n \rightarrow \infty} \underline{\pi}(n, \gamma)$  exists then  $\lim_{n \rightarrow \infty} \underline{\pi}(n, \gamma)/C(n, \gamma)$  exists and satisfies equation (4.2.40a) and (4.2.40b). Necessary conditions for the convergence of  $\underline{\pi}(n, \gamma)$  have not yet been found; a sufficient condition is given by the following theorem.

**Theorem 4.2.2** Let  $\tilde{P}$  have the prior distribution function  $F(\underline{P} | \gamma) \in \mathcal{F}$ , a family of distributions closed under consecutive sampling which is continuous on the boundary of  $\mathcal{S}_N$ . Let  $\underline{\pi}(n, \gamma)$  be defined by equation (4.2.42) with the constant terminal functions

$$\bar{\pi}_i(0, \gamma) = p_i, \quad i=1, \dots, N \quad (\text{4.2.44})$$

$$\gamma \in \Psi$$

where  $p = (p_1, \dots, p_N)$  is a stochastic vector. Then, for  $i=1, \dots, N$ ,  $\lim_{n \rightarrow \infty} \bar{\pi}_i(n, \gamma)$  exists and is equal to  $E[\tilde{\pi}_i | \gamma]$ . Moreover,

$$\sum_{i=1}^N \lim_{n \rightarrow \infty} \bar{\pi}_i(n, \gamma) = 1. \quad \gamma \in \Psi \quad (\text{4.2.45})$$

Thus,  $\lim_{n \rightarrow \infty} \underline{\pi}(n, \gamma)$  satisfies equation (4.2.40).

**Proof.** The theorem is proved by showing that

$$\bar{\pi}_j(n, \gamma) = \sum_{i=1}^N p_i \tilde{p}_{ij}^{(n)}(\gamma), \quad j=1, \dots, N \quad (\text{4.2.46})$$

$$n=1, 2, 3, \dots$$

$$\gamma \in \Psi$$



from which it follows, by Theorem 4.2.5, that

$$\begin{aligned}\lim_{n \rightarrow \infty} \bar{\pi}_j(n, \tau) &= \sum_{k=1}^N p_k E[\tilde{\pi}_j(\tau)] \\ &= E[\tilde{\pi}_j(\tau)]. \quad j=1, \dots, N \quad (4.2.47)\end{aligned}$$

Equation (4.2.46) is established inductively. For  $n=1$ ,  $\tilde{p}_{1j}^{(n)}(\tau) = \tilde{p}_{1j}(\tau)$  and (4.2.46) holds. Assume it is true for  $n$ . Then, using equation (4.1.2a)

$$\begin{aligned}\bar{\pi}_j(n+1, \tau) &= \sum_{k=1}^N \bar{\pi}_k(n, T_{kj}(\tau)) \tilde{p}_{kj}(\tau) \\ &= \sum_{k=1}^N \sum_{i=1}^N p_i \tilde{p}_{ik}^{(n)}(T_{kj}(\tau)) \tilde{p}_{kj}(\tau) \\ &= \sum_{i=1}^N p_i \tilde{p}_{ij}^{(n+1)}(\tau), \quad (4.2.48)\end{aligned}$$

proving the assertion. Summing (4.2.46) over  $j$ , we have  $\sum_{j=1}^N \bar{\pi}_j(n, \tau) = 1$  ( $n=1, 2, \dots$ ;  $\tau \in \mathbb{T}$ ), since

$$\sum_{j=1}^N \tilde{p}_{1j}^{(n)}(\tau) = \int \sum_{j=1}^N \tilde{p}_{1j}^{(n)} dF_P(\tau) = 1. \quad (4.2.49)$$

$\mathcal{B}_N$

$j=1, \dots, N$   
 $n=1, 2, \dots$   
 $\tau \in \mathbb{T}$

Q.E.D.

By letting  $p_i = \delta_{ik}$  for some fixed index  $k$ , equation (4.2.46) becomes

$$\bar{\pi}_j(n, \tau) = \tilde{p}_{kj}^{(n)}(\tau), \quad \begin{matrix} j=1, \dots, N \\ n=1, 2, 3, \dots \\ \tau \in \mathbb{T} \end{matrix} \quad (4.2.50)$$

and it is seen that the approximation of  $\bar{\pi}_j(\tau)$  by  $\tilde{p}_{kj}^{(n)}(\tau)$  is a special case of the method of successive approximations defined by (4.2.42) and (4.2.44).

Another approximation of interest is based upon (4.2.42) and the terminal function  $\underline{\pi}(0, \tau)$  defined by



$$\bar{\pi}_i(0, \gamma) = \pi_i(\bar{P}(\gamma)), \quad i=1, \dots, N \quad (4.2.51)$$
$$\gamma \in \Sigma$$

where  $\underline{\pi}(\bar{P}(\gamma))$  is the steady-state probability vector corresponding to  $\bar{P}(\gamma)$ , the mean of the distribution function  $P(P|\gamma)$ . We have not been able to prove convergence of  $\underline{\pi}(n, \gamma)$  in this case, but limited computational experience with this approximation, using a matrix beta prior distribution, suggests that convergence does occur and, in some cases, is more rapid than the convergence of  $\underline{p}_{ij}^{(n)}(\gamma)$ .

Some numerical results based on the recursive programming of (4.2.42) with the terminal functions (4.2.51) are displayed in Tables 4.2.4-4.2.6. A matrix beta prior distribution was used in all cases. The program is given in Appendix D.

A two state process is considered in Table 4.2.4. The transition matrix has the same prior distribution as was used to compute Table 4.2.1, where it was seen that  $\underline{\pi}(\underline{m}) = (0.93293 \quad 0.06707)$ . In Table 4.2.4 the approximation  $\underline{\pi}(n, \underline{m})$  defined by (4.2.42) is given in column two and the normalizing constant,  $C(n, \underline{m}) = \bar{\pi}_1(n, \underline{m}) + \bar{\pi}_2(n, \underline{m})$  is given in column three. In column four it is seen that

$\lim_{n \rightarrow \infty} \frac{1}{C(n, \underline{m})} \underline{\pi}(n, \underline{m}) = \underline{\pi}(\underline{m})$ , with three-place accuracy on the first iteration and four-place accuracy on the second iteration. The eight entries of Table 4.2.4 required 0.62 minutes of computation time on an IBM 7094 machine.

In Table 4.2.5 a  $2 \times 2$  transition matrix is treated which has the same prior distribution as the matrix considered in Table 4.2.2. This is a relatively loose prior distribution and it is seen that convergence of  $\frac{1}{C(n, \underline{m})} \underline{\pi}(n, \underline{m})$  is slow, although comparison with Table 4.2.2 indicates that this approximation has smaller error than the approximation



$n$	$\bar{\pi}(n, \underline{m})$	$C(n, \underline{m})$	$\frac{1}{C(n, \underline{m})} \bar{\pi}(n, \underline{m})$
1	(0.93301 0.06724)	1.00025	(0.93278 0.06722)
2	(0.93311 0.06713)	1.00024	(0.93289 0.06711)
3	(0.93313 0.06710)	1.00023	(0.93292 0.06708)
4	(0.93311 0.06709)	1.00020	(0.93292 0.06708)
5	(0.93310 0.06709)	1.00019	(0.93292 0.06708)
6	(0.93308 0.06709)	1.00017	(0.93292 0.06708)
7	(0.93307 0.06708)	1.00015	(0.93293 0.06707)
8	(0.93306 0.06708)	1.00014	(0.93293 0.06707)

$$\underline{m} = \begin{bmatrix} 14.105 & 0.988 \\ 21.313 & 3.367 \end{bmatrix} \quad \bar{P} = \begin{bmatrix} 0.93454 & 0.06546 \\ 0.86357 & 0.13643 \end{bmatrix}$$

$$V(\bar{P}) = \begin{bmatrix} 0.0038 & 0.0038 \\ 0.0049 & 0.0048 \end{bmatrix}$$

Computation Time: 0.62 minutes.

Table 4.2.4



$n$	$\bar{U}(n, \underline{m})$	$C(n, \underline{m})$	$\frac{1}{C(n, \underline{m})}$	$\bar{U}(n, \underline{\underline{m}})$
1	(0.43377 0.63815)	1.07192	(0.40467 0.59533)	
2	(0.44423 0.62777)	1.07200	(0.41439 0.58561)	
3	(0.44447 0.62433)	1.06880	(0.41586 0.58414)	
4	(0.44443 0.61940)	1.06383	(0.41776 0.58224)	
5	(0.44308 0.61632)	1.05940	(0.41824 0.58176)	
6	(0.44211 0.61317)	1.05528	(0.41895 0.58105)	
7	(0.44085 0.61082)	1.05170	(0.41921 0.58079)	
8	(0.43992 0.60858)	1.04850	(0.41957 0.58043)	
9	(0.43890 0.60679)	1.04569	(0.41972 0.58028)	
10	(0.43808 0.60511)	1.04319	(0.41994 0.58006)	
11	(0.43725 0.60370)	1.04095	(0.42005 0.57995)	
12	(0.43656 0.60230)	1.03894	(0.42020 0.57980)	
13	(0.43583 0.60124)	1.03712	(0.42028 0.57972)	
14	(0.43529 0.60018)	1.03547	(0.42038 0.57962)	

$$\underline{m} = \begin{bmatrix} 0.251 & 0.352 \\ 0.617 & 1.120 \end{bmatrix}$$

$$\underline{\underline{m}} = \begin{bmatrix} 0.41625 & 0.58375 \\ 0.35521 & 0.64479 \end{bmatrix}$$

$$v(\tilde{p}) = \begin{bmatrix} 0.1516 & 0.1516 \\ 0.0837 & 0.0837 \end{bmatrix}$$

Computation Time: 5.03 minutes ( $n = 1, \dots, 12$ )  
 15.03 minutes ( $n = 13, 14$ ).



$n$	$\bar{P}(n, \underline{m})$			$C(n, \underline{m})$	$\frac{1}{C(n, \underline{m})} \bar{P}(n, \underline{m})$
1	(0.32341	0.37234	0.30058)	1.00193	(0.32797 0.37185 0.30018)
2	(0.32915	0.37145	0.30142)	1.00202	(0.32899 0.37070 0.30081)
3	(0.32969	0.37185	0.30100)	1.00248	(0.32881 0.37093 0.30026)
4	(0.32990	0.37160	0.30123)	1.00273	(0.32900 0.37059 0.30041)
5	(0.33007	0.37171	0.30110)	1.00288	(0.32912 0.37064 0.30034)
6	(0.33017	0.37162	0.30117)	1.00296	(0.32920 0.37052 0.30028)
7	(0.33023	0.37164	0.30112)	1.00299	(0.32924 0.37054 0.30022)
8	(0.33025	0.37160	0.30119)	1.00298	(0.32927 0.37050 0.30023)
9	(0.33026	0.37160	0.30110)	1.00296	(0.32929 0.37050 0.30021)

$$\underline{m} = \begin{bmatrix} 18.265 & 2.102 & 3.910 \\ 2.385 & 5.168 & 10.111 \\ 1.005 & 7.612 & 1.212 \end{bmatrix}$$

$$\bar{P} = \begin{bmatrix} 0.75236 & 0.08653 & 0.16106 \\ 0.13502 & 0.29257 & 0.57241 \\ 0.10225 & 0.77444 & 0.12331 \end{bmatrix}$$

$$V(\bar{P}) = \begin{bmatrix} 0.0074 & 0.0031 & 0.0053 \\ 0.0063 & 0.0111 & 0.0131 \\ 0.0085 & 0.0161 & 0.0100 \end{bmatrix}$$

Computation Time: 2.15 minutes ( $n = 1, \dots, ?$ ).



has smaller error than the approximation  $\hat{p}_{ij}^{(n)}(\underline{\eta}_1)$ . The first twelve iterations required a total of 5.03 minutes, while iteration thirteen and fourteen consumed 15.03 minutes, illustrating the exponential growth of the computation time with  $n$ .

A  $3 \times 3$  transition matrix which has the same prior distribution as was used in computing Table 4.2.3 is considered in Table 4.2.6. In this case,  $\frac{1}{C(n, \underline{\eta}_1)} \tilde{\Pi}(n, \underline{\eta}_1)$  converges faster than the approximation  $\hat{p}_{ij}^{(n)}(\underline{\eta}_1)$ . Two-place accuracy is achieved on the first iteration, with three-place accuracy on the third iteration. The computation time for the first seven entries was 2.15 minutes; the time for entries 8 and 9 is not available.

#### 4.2.4 Variance-Covariance Matrix of $\tilde{\Pi}$ . Let

$$\bar{\pi}_{ij}(\psi) = \int_{\mathcal{S}_N} \pi_i(\underline{p}) \pi_j(\underline{p}) dF(\underline{p} | \psi) \quad i, j = 1, \dots, N \quad (4.2.52)$$

be the expected value of  $\tilde{\pi}_i \tilde{\pi}_j$  when  $\tilde{\underline{p}}$  has the distribution function  $F(\underline{p} | \psi)$ . If  $F(\underline{p} | \psi)$  is continuous on the boundary of  $\mathcal{S}_N$ , Theorem 4.2.2 implies the existence of the integral (4.2.52). If  $F(\underline{p} | \psi) \in \mathcal{F}$ , a family of distributions continuous in  $\psi$ , then Theorem 2.4.3 implies that  $\bar{\pi}_{ij}(\psi)$  is a continuous function of  $\psi$ . When  $\mathcal{F}$  is also closed under the consecutive sampling rule, the following theorem shows that  $\bar{\pi}_{ij}(\psi)$  can be approximated by  $E[\hat{p}_{ij}^{(n)}(\underline{p}) | \psi]$ .

**Theorem 4.2.8** If the prior distribution function of  $\tilde{\underline{p}}$  is  $F(\underline{p} | \psi) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under consecutive sampling, then



$$\lim_{\substack{n \rightarrow \infty \\ v \rightarrow \infty}} E[\tilde{p}_{\alpha i}^{(n)} \tilde{p}_{\beta j}^{(v)} | \gamma] = \bar{\pi}_{ij}(\gamma). \quad (4.2.53)$$

$\alpha, \beta, i, j = 1, \dots, N$   
 $\gamma \in \Xi$

Proof. Let  $\epsilon > 0$  be given. By Lemma 4.2.4,  $\{p_{\alpha i}^{(n)} p_{\beta j}^{(v)}\}_{\gamma \in \Xi} \rightarrow \pi_i(P) \pi_j(P)$  uniformly on  $\mathcal{S}_N^{\alpha}$  for any  $\alpha \in (0, 1)$ . Arguing as in the proof of Theorem 4.2.5, we may choose  $n$  and  $v$  sufficiently large and  $\alpha > 0$  sufficiently small that

$$\left| E[\tilde{p}_{\alpha i}^{(n)} \tilde{p}_{\beta j}^{(v)} | \gamma] - \bar{\pi}_{ij}(\gamma) \right| < \epsilon, \quad (4.2.54)$$

proving the theorem. Q.E.D.

Theorem 4.2.9 If the prior distribution function of  $\tilde{P}$  is  $F(P | \gamma) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under consecutive sampling, then the product moment  $\bar{\pi}_{ij}(\gamma)$  satisfies the following functional equations:

$$\bar{\pi}_{ij}(\gamma) = \sum_{k=1}^N \sum_{m=1}^N \tilde{p}_{ki}(\gamma) \tilde{p}_{mj}(T_{ki}(\gamma)) \bar{\pi}_{km}(T_{mj}(T_{ki}(\gamma))), \quad (4.2.55a)$$

$i, j = 1, \dots, N$   
 $\gamma \in \Xi$

$$\sum_{i=1}^N \sum_{j=1}^N \bar{\pi}_{ij}(\gamma) = 1. \quad \gamma \in \Xi \quad (4.2.55b)$$

Remark. The condition (4.2.55b) is necessary to insure a unique solution to the functional equation (4.2.55a). Sufficient conditions for a unique solution have not yet been found.

Proof. Since

$$\pi_i(P) \pi_j(P) = \sum_{k=1}^N \sum_{m=1}^N \pi_k(P) \pi_m(P) p_{ki} p_{mj}, \quad P \in \mathcal{S}_N^{\alpha} \quad (4.2.56)$$



equation (4.2.55a) follows from Lemma 2.3.2. Equation (4.2.55b) follows by summing (4.2.52) over i and j. Q.E.D.

Given  $\bar{\pi}_{ij}(\tau)$ ,  $\bar{\pi}_i(\tau)$ , and  $\bar{\pi}_j(\tau)$ , the covariance between  $\tilde{\pi}_i$  and  $\tilde{\pi}_j$  is computed from the relation

$$\text{cov} [\tilde{\pi}_i, \tilde{\pi}_j | \tau] = \bar{\pi}_{ij}(\tau) - \bar{\pi}_i(\tau) \bar{\pi}_j(\tau). \quad (4.2.57)$$

$i, j = 1, \dots, N$   
 $\tau \in \mathbb{T}$

#### 4.3 Expected Discounted Reward Vector.

Consider a Markov chain which is operated indefinitely under a fixed policy with initial state i. If  $\tilde{P} = P$ , let  $V_i(P)$  be the conditional expectation of the total discounted reward earned over an infinite period when the chain starts in state i and let  $V(P) = (V_1(P), \dots, V_N(P))$  be the vector of expected discounted rewards. Howard\* has shown that, for any  $P \in \mathcal{P}_N$ , including periodic and multiple-chain transition matrices,

$$V_i(P) = \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{ij}^{(n)} p_{jk} P_{jk}. \quad (4.3.1)$$

$i = 1, \dots, N$

$P \in \mathcal{P}_N$

$0 \leq \beta < 1$

When  $\tilde{P}$  is a random matrix  $V(\tilde{P})$  is a random vector. In this section the mean and the variance-covariance matrix of  $V(\tilde{P})$  are studied and expressions for elements of these unconditional expectations are found. A set of functional equations for the expected value of  $V_i(\tilde{P})$  is derived which is closely related to equation (3.1.5), which was discussed in connection with the discounted adaptive control problem. This relation is

\* [2], p.82.



used to obtain a method of successive approximations for the numerical calculation of the expected value of  $V_{\frac{1}{n}}(\tilde{P})$ . Some examples are presented in Section 4.3.3.

**4.3.1 Expected Value of  $V_{\frac{1}{n}}(\tilde{P})$ .** Let  $\tilde{P}$  have the prior distribution function  $F(P|\gamma)$  and let

$$\tilde{V}_{\frac{1}{n}}(\gamma) = \int_{\mathcal{S}_N} V_{\frac{1}{n}}(P)dF(P|\gamma) \quad i=1, \dots, N \quad (4.3.2)$$

be the expected value of  $V_{\frac{1}{n}}(\tilde{P})$ . The first theorem shows that this expectation exists and provides a formula for  $\tilde{V}_{\frac{1}{n}}(\gamma)$  in terms of the expected n-step transition probabilities,  $\tilde{p}_{ij}^{(n)}(\gamma)$ , when the prior distribution belongs to a family closed under consecutive sampling. A preliminary lemma is required.

**Lemma 4.3.1.** If  $V_{\frac{1}{n}}(P)$  is defined by the infinite series (4.3.1) with  $0 \leq \beta < 1$ , then the series converges uniformly in  $P$  and  $V_{\frac{1}{n}}(P)$  is continuous on  $\mathcal{S}_N$  ( $i=1, \dots, N$ ).

**Proof.** For any finite  $n$ , the functions

$$w_i(n, P) = \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{ij}^{(n)} p_{jk} r_{jk}, \quad (4.3.3)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $P \in \mathcal{S}_N$

are continuous functions of  $P$  on  $\mathcal{S}_N$ . Moreover, if

$$R^* = \max_{1,j} \left\{ |r_{1j}| \right\}, \quad (4.3.4)$$

the functions  $w_i(n, P)$  are bounded on  $\mathcal{S}_N$ .

$$|w_i(n, P)| \leq R^* \beta^n. \quad (4.3.5)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $P \in \mathcal{S}_N$



Since

$$\sum_{n=0}^{\infty} R^n \beta^n = \frac{\beta^n}{1-\beta} < \infty \quad (4.3.6)$$

for  $0 \leq \beta < 1$ ,  $\sum_{n=0}^{\infty} v_1(n, P)$  converges uniformly\* to  $\tilde{V}_1(P)$  on  $\mathcal{S}_N$  and  $\tilde{V}_1(P)$  is a continuous function of  $P$  on  $\mathcal{S}_N$ . Q.E.D.

Theorem 4.3.2 Let  $\tilde{P}$  have the prior distribution function  $F(P|\psi)$ . Then the expectation  $\tilde{V}_1(\psi)$  defined by equation (4.3.2) exists. If  $F(P|\psi) \in \mathcal{F}$ , a family of distributions closed under the consecutive sampling rule, then

$$\tilde{V}_1(\psi) = \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{1j}^{(n)} (T_{jk}(\psi)) \tilde{p}_{jk}(\psi) r_{jk}, \quad (4.3.7)$$

$j=1, \dots, N$   
 $\psi \in \mathcal{F}$   
 $0 \leq \beta < 1$

where  $p_{1j}^{(n)}(T_{jk}(\psi))$  is defined by equation (4.1.1)

Proof. The integral (4.3.2) exists by virtue of the continuity of the bounded function  $V_1(P)$  on  $\mathcal{S}_N$ . Since the infinite series (4.3.1) converges uniformly,

$$\tilde{V}_1(\psi) = \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N r_{jk} \int_{\mathcal{S}_N} p_{1j}^{(n)} p_{jk} dF(P|\psi). \quad (4.3.8)$$

If  $\mathcal{F}$  is closed under consecutive sampling, Lemma 2.3.2 yields (4.3.7). Q.E.D.

In Section 4.4.2 we shall discuss approximations to  $\tilde{V}_1(\psi)$  which are based on equation (4.3.7). The results of the following paragraph provide a different basis for computation of  $\tilde{V}_1(\psi)$ .

#### 4.3.2 A Functional Equation for $\tilde{V}_1(\psi)$ .

We now relate  $\tilde{V}_1(\psi)$  to

---

\* Rudin [35], p. 134.



$v_1(\psi)$ , the maximum expected discounted reward discussed in connection with the adaptive control problem.

**Theorem 4.3.3** If  $\tilde{P}$  has the prior distribution function  $F(P| \psi) \in \mathcal{F}$ , a family of distributions closed under consecutive sampling, then  $\tilde{V}(\psi) = (\tilde{v}_1(\psi), \dots, \tilde{v}_N(\psi))$  satisfies the following set of simultaneous functional equations,

$$\tilde{v}_i(\psi) = \tilde{q}_i(\psi) + \beta \sum_{j=1}^N \tilde{p}_{ij}(\psi) \tilde{v}_j(T_{ij}(\psi)), \quad (4.3.9)$$

$i=1, \dots, N$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

where  $\tilde{q}_i(\psi)$ , the expected one-step transition reward when in state  $i$ , is defined by equation (3.1.4).

**Proof.** Letting

$$q_i(P) = \sum_{k=1}^N p_{ik} r_{ik}, \quad i=1, \dots, N \quad (4.3.10)$$

$P \in \mathcal{S}_N$

equation (4.3.1) can be written

$$\begin{aligned}
 V_i(P) &= \sum_{k=1}^N p_{ik} r_{ik} + \beta \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{ij} p_{jk} r_{jk}^{(n+1)} \\
 &= q_i(P) + \beta \sum_{m=1}^N p_{im} \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{mj}^{(n)} p_{jk} r_{jk} \\
 &= q_i(P) + \beta \sum_{m=1}^N p_{im} V_m(P). \quad i=1, \dots, N \quad (4.3.11)
 \end{aligned}$$

$P \in \mathcal{S}_N$

Thus, changing the index of summation and using Lemma 2.3.2,

$$\begin{aligned}
 \tilde{v}_i(\psi) &= \tilde{q}_i(\psi) + \beta \sum_{j=1}^N \int_{\mathcal{S}_N} p_{ij} V_j(P) dF(P|\psi) \\
 &= \tilde{q}_i(\psi) + \beta \sum_{j=1}^N \tilde{p}_{ij}(\psi) \tilde{v}_j(T_{ij}(\psi)). \quad (4.3.12)
 \end{aligned}$$

Q.E.D.



Equation (4.3.9) has the same form as equation (3.1.5) and may be interpreted as a discounted adaptive control equation in which there is exactly one alternative in each state. The results of Sections 3.2 and 3.3 apply and are summarized in the following theorem which is stated without proof, since the proofs in the more general case of many alternatives in each state are given in Chapter 3.

Theorem 4.3.4 There exists a unique bounded vector function,  $\tilde{v}(\psi) = (\tilde{v}_1(\psi), \dots, \tilde{v}_N(\psi))$ , which satisfies equation (4.3.9). Let the sequences of functions  $\{\tilde{v}_i(n, \psi)\}$ ,  $i=1, \dots, N$ , be defined by the equations

$$\tilde{v}_i(n+1, \psi) = \tilde{q}_i(\psi) + \beta \sum_{j=1}^N \tilde{p}_{ij}(\psi) \tilde{v}_j(n, T_{ij}(\psi)), \quad (4.3.13a)$$

$i=1, \dots, N$

$n=0, 1, 2, \dots$

$\psi \in \Psi$

$0 \leq \beta < 1$

$i=1, \dots, N$

$\psi \in \Psi$

$$\tilde{v}_i(0, \psi) = v_i(\psi). \quad (4.3.13b)$$

Then, provided the terminal functions  $v_i(\psi)$  are bounded,

$$|v_i(\psi)| \leq v \quad i=1, \dots, N \quad (4.3.14)$$

the sequence  $\{\tilde{v}_i(n, \psi)\}$  converges uniformly to  $\tilde{v}_i(\psi)$ , the unique bounded solution of (4.3.9). If the terminal functions are constants,

$$v_i(\psi) = c_i \quad i=1, \dots, N \quad (4.3.15)$$

and if  $v^* = \max_i \{v_i\}$ ,  $v^* = \min_i \{v_i\}$ ,  $R = \max_{i,j} \{r_{ij}\}$  and  $r = \min_{i,j} \{r_{ij}\}$ ,

then the error of the  $n$ th approximant,

$$\tilde{e}_i(n, \psi) = \tilde{v}_i(\psi) - \tilde{v}_i(n, \psi), \quad (4.3.16)$$

has the bound



$$|\bar{v}_1(n, \psi)| \leq \beta^n \left[ \max \left\{ \frac{R}{1-\beta} - v^*, v^* - \frac{R}{1-\beta} \right\} \right]. \quad (4.3.17)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

If, furthermore,

$$v^* - \beta v^* \leq r, \quad (4.3.18)$$

then  $\{\bar{v}_1(n, \psi)\}$  is a monotone increasing sequence with limit  $\bar{v}_1(\psi)$  and the bound (4.3.17) becomes

$$0 \leq \bar{v}_1(n, \psi) \leq \beta^n \left( \frac{R}{1-\beta} - v^* \right). \quad (4.3.19)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

If

$$v^* - \beta v^* \geq R \quad (4.3.20)$$

then  $\{\bar{v}_1(n, \psi)\}$  is a monotone decreasing sequence with limit  $\bar{v}_1(\psi)$  and

$$\beta^n \left[ \frac{R}{1-\beta} - v^* \right] \leq \bar{v}_1(n, \psi) \leq 0. \quad (4.3.21)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\psi \in \Psi$   
 $0 \leq \beta < 1$

**4.3.3 Numerical Example.** To illustrate the above remarks, some sample computations based on equation (4.3.13) are displayed in Tables 4.3.1-4.3.4\*. These tables contain values of

$$\bar{v}(n, \underline{\psi}) = \begin{bmatrix} \bar{v}_1(n, \underline{\psi}) \\ \bar{v}_2(n, \underline{\psi}) \end{bmatrix}$$

under various policies in a two-state Markov chain with two alternatives in each state when  $\tilde{P}$  has a matrix beta distribution. The discount factor is  $\beta = 0.2$ . The reward matrix and the prior distribution are the same as

---

\* See Appendix E for the program listing.



$n$	$\hat{V}(n, \underline{m})$	$\hat{\pi}(n, \underline{m})$	$\Delta(n)$
0	9.660 5.596	0.334 ~0.052	11.904
1	9.663 5.596	0.331 ~0.052	2.381
2	9.936 5.544	0.058 ~0.005	0.476
3	9.982 5.544	0.012 0.000	0.095
4	9.991 5.544	0.003 0.000	0.019
5	9.993 5.544	0.001 0.000	0.004
6	9.993 5.544	0.001 0.000	0.001
7	9.994 5.544	0.000 0.000	0.000
8	9.994 5.544	0.000 0.000	0.000
9	9.994 5.544	0.000 0.000	0.000

Policy: (1,1).

$$\hat{V}(0, \underline{m}) = \begin{bmatrix} 9.660 \\ 5.596 \end{bmatrix}$$

Computation time: 0.56 minutes.

$\beta = 0.2$ .



$n$	$\bar{V}(n, \underline{m})$	$\bar{\delta}(n, \underline{m})$	$\Delta(n)$
0	10.253	0.265	9.747
	13.278	0.390	
1	10.251	0.264	1.949
	13.276	0.392	
2	10.470	0.048	0.390
	13.586	0.082	
3	10.507	0.011	0.078
	13.652	0.016	
4	10.516	0.002	0.016
	13.665	0.003	
5	10.517	0.001	0.003
	13.667	0.001	
6	10.518	0.000	0.001
	13.668	0.000	
7	10.518	0.000	0.000
	13.668	0.000	
8	10.518	0.000	0.000
	13.668	0.000	
9	10.518	0.000	0.000
	13.668	0.000	

Policy : (1,2).

$$\bar{V}(0, \underline{m}) = \begin{bmatrix} 10.253 \\ 13.278 \end{bmatrix}$$

Computation time: 0.39 minutes.

$\beta = 0.2$ .



$n$	$\tilde{Y}(n, \underline{m})$	$\tilde{g}(n, \underline{m})$	$\Delta(n)$
0	5.616	-0.043	12.027
	5.473	-0.059	
1	5.616	-0.043	2.405
	5.473	-0.059	
2	5.595	-0.012	0.481
	5.429	-0.015	
3	5.576	-0.003	0.096
	5.417	-0.003	
4	5.574	-0.001	0.019
	5.414	0.000	
5	5.573	0.000	0.004
	5.414	0.000	
6	5.573	0.000	0.001
	5.414	0.000	
7	5.573	0.000	0.000
	5.414	0.000	
8	5.573	0.000	0.000
	5.414	0.000	
9	5.573	0.000	0.000
	5.414	0.000	

Policy: (2,1).

$$\tilde{y}(0, \underline{m}) = \begin{bmatrix} 5.616 \\ 5.473 \end{bmatrix}$$

Computation time: 0.55 minutes.

$\beta = 0.2$ .



$n$	$\tilde{v}(n, \underline{m})$	$\tilde{\sigma}(n, \underline{m})$	$\Delta(n)$
0	6.042	0.042	13.958
	12.708	0.389	
1	6.042	0.042	2.792
	12.707	0.390	
2	5.988	-0.012	0.558
	13.043	0.054	
3	5.999	-0.001	0.112
	13.084	0.013	
4	6.000	0.000	0.022
	13.095	0.002	
5	6.000	0.000	0.004
	13.097	0.000	
6	6.000	0.000	0.001
	13.097	0.000	
7	6.000	0.000	0.000
	13.097	0.000	
8	6.000	0.000	0.000
	13.097	0.000	
9	6.000	0.000	0.000
	13.097	0.000	

Policy: (2,2).

$$\tilde{v}(0, \underline{m}) = \begin{bmatrix} 6.042 \\ 12.708 \end{bmatrix}$$

Computation time: 0.62 minutes.

$\beta = 0.2$ .



those used in computing Tables 3.5.1 - 3.5.3 for the adaptive control problem (see equations (3.5.1) - (3.5.4)). For each of the four possible policies,  $\underline{\Sigma}$ , the terminal functions are given by

$$\tilde{V}_1(0, \underline{M}) = V_1(\tilde{P}(\sigma)), \quad i=1,2$$

the expected discounted reward starting from state 1 when  $\tilde{P}(\sigma) = \tilde{P}(\underline{\Sigma})$ .

It is seen that convergence takes place, in each case, by the seventh iteration when accuracy to the third decimal place is desired. The computation time indicated at the end of each table is the total time required to calculate all nine iterations on an IBM 7094 computer.

Also displayed in each table is the error vector,

$$\bar{e}(n, \underline{M}) = \begin{bmatrix} \tilde{V}_1(n, \underline{M}) \\ \tilde{V}_2(n, \underline{M}) \end{bmatrix} - \tilde{Y}(n, \underline{M}),$$

taking  $\tilde{Y}(\underline{M}) = \tilde{Y}(0, \underline{M})$ . The last column of each table contains values of

$$\Delta(n) = \beta^n [\max \left\{ \frac{P}{1-\beta} - v^*, v^* - \frac{k}{1-\beta} \right\}],$$

the absolute error bound of equation (4.3.17).

Comparing these calculations with those of Tables 3.5.1 - 3.5.3, it is seen that, in this example, the adaptive control problem and the problem of choosing a terminal policy which maximizes  $V_1(\underline{M})$  both have the same optimal initial policy and the same total expected reward.

**4.3.4 Variance-Covariance Matrix.** We conclude this section with a formula for the covariance of  $V_1(\tilde{P})$  and  $V_2(\tilde{P})$ . This equation involves terms of the form  $E[V_{ij}^{(n)} | \Gamma]$ , which can be computed with the aid of Theorem 4.1.3. Approximations to  $\text{cov}[V_1(\tilde{P}), V_2(\tilde{P}) | \Gamma]$  are considered in Section 4.4.2.



**Theorem 4.3.5** If  $\tilde{P}$  has the distribution function  $F(\tilde{P} | \gamma) \in \mathcal{F}_\gamma$ , a family of distributions closed under consecutive sampling, then the covariance between  $V_i(\tilde{P})$  and  $V_j(\tilde{P})$  is given by

$$\text{cov}[V_i(\tilde{P}), V_j(\tilde{P}) | \gamma] = \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} \beta^{n+v} \sum_{a, v, k, m=1}^N \tilde{p}_{ak}^{(n)} r_{ak} r_{mv} \\ E[\tilde{p}_{1a}^{(n)} \tilde{p}_{jm}^{(v)} | T_{mv}(T_{ak}(\gamma))] \tilde{p}_{mv}^{(n)}(T_{ak}(\gamma)) = \tilde{p}_{1a}^{(n)}(T_{ak}(\gamma)) \tilde{p}_{jm}^{(v)}(T_{mv}(\gamma)) \tilde{p}_{mv}^{(n)}(\gamma) \\ i, j=1, \dots, N \quad (4.3.22) \\ \gamma \in \mathbb{P} \\ 0 \leq \beta < 1$$

**Proof.** The expected value of the product  $V_i(\tilde{P}) V_j(\tilde{P})$  is

$$E[V_i(\tilde{P}) V_j(\tilde{P}) | \gamma] = \int_{\mathcal{S}_N} \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} w_i(n, P) w_j(v, P) dF(\tilde{P} | \gamma), \quad (4.3.23)$$

where

$$w_i(n, P) = \beta^n \sum_{c=1}^N \sum_{k=1}^N p_{1c}^{(n)} p_{ck}^{(n)} r_{ak}. \quad i=1, \dots, N \quad (4.3.24) \\ n=0, 1, 2, \dots \\ P \in \mathcal{S}_N$$

If  $R^* = \max_{i,j} \{|r_{ij}| \}$ , then

$$|w_i(n, P) w_j(v, P)| \leq (R^*)^2 \beta^{n+v}. \quad i, j=1, \dots, N \quad (4.3.25) \\ n, v=0, 1, 2, \dots \\ P \in \mathcal{S}_N$$

Since

$$(R^*)^2 \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} \beta^{n+v} = \frac{(R^*)^2}{(1-\beta)^2} < \infty, \quad 0 \leq \beta < 1 \quad (4.3.26)$$

the double sum in the integrand of (4.3.23) converges uniformly\* to  $V_i(\tilde{P}) V_j(\tilde{P})$  on  $\mathcal{S}_N$ . Thus,

$$E[V_i(\tilde{P}) V_j(\tilde{P}) | \gamma] = \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} \beta^{n+v} \sum_{a, v, k, m=1}^N r_{ak} r_{mv} \int_{\mathcal{S}_N} p_{1a}^{(n)} p_{ak}^{(n)} p_{jm}^{(v)} p_{mv}^{(v)} dF(\tilde{P} | \gamma) \quad (4.3.27)$$

\* Radin [35], p. 134.



Applying Lemma 2.3.2 twice,

$$E [v_i(\tilde{P})v_j(\tilde{P}) | \gamma] =$$

$$\sum_{n=0}^{\infty} \sum_{v=0}^{\infty} \beta^{n+v} \sum_{a, v, k, m=1}^N \bar{p}_{ak}(\gamma) \bar{p}_{av}(\tau_{ak}(\gamma)) r_{ak} r_{av} E[p_{ak}^{(n)} p_{jk}^{(v)} | \tau_{av}(\tau_{ak}(\gamma))]. \quad (4.3.28)$$

Subtracting

$$\tilde{v}_i(\gamma) \tilde{v}_j(\gamma) = \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} \beta^{n+v} \sum_{a, v, k, m=1}^N \bar{p}_{ak}(\gamma) \bar{p}_{av}(\gamma) r_{ak} r_{av} \frac{p_{jk}^{(v)}}{p_{jk}^{(n)}} (\tau_{ak}(\gamma)) \tilde{p}_{jk}^{(v)} \tau_{av} \quad (4.3.29)$$

from (4.3.28), equation (4.3.22) is obtained. Q.E.D.

**Theorem 4.3.6** If  $F(P|\gamma) \in \mathcal{F}_1$ , a family of distributions continuous in  $\gamma$ , then  $\tilde{v}_i(\gamma)$  and  $\text{cov}[v_i(\tilde{P}), v_j(\tilde{P}) | \gamma]$  are continuous functions of  $\gamma$  on  $\mathbb{S}_N$  ( $i, j=1, \dots, N$ ).

**Proof.** Lemma 4.3.1 implies that the bounded function  $v_i(\tilde{P})$  is integrable on  $\mathbb{S}_N$ . Thus, by Theorem 2.4.3,  $\tilde{v}_i(\gamma)$  and  $E[v_i(\tilde{P})v_j(\tilde{P}) | \gamma]$  are continuous on  $\mathbb{S}$ . Q.E.D.

#### 4.4 The Process Gain.

Consider a Markov chain operating indefinitely under a fixed policy. The conditional expected reward per transition, given that  $\tilde{P} = P$ , an ergodic transition matrix, is

$$g(P) = \sum_{i=1}^N \sum_{j=1}^N \pi_i(P) p_{ij} r_{ij}, \quad (4.4.1)$$

and is known as the gain of the process. When  $\tilde{P}$  is a random matrix with the distribution function  $F(P|\gamma)$ , which is assumed to be continuous on the boundary of  $\mathbb{S}_N$ , then  $g(\tilde{P})$  is a random variable. The mean and variance of  $g(\tilde{P})$  are investigated in this section, assuming that  $F(P|\gamma)$  belongs



to a family of distributions closed under the consecutive sampling rule.

**4.4.1 Mean and Variance of  $g(\tilde{P})$ .** Let the expected value of  $g(\tilde{P})$  be

$$\bar{g}(\psi) = \int_{\mathcal{S}_N} g(P)dF(P|\psi). \quad (4.4.2)$$

Equation (4.4.1) shows that  $g(P)$  is continuous and bounded, hence, integrable on  $\mathcal{S}_N^*$ . If  $F(P|\psi)$  is continuous on the boundary of  $\mathcal{S}_N$ , then Lemma 4.2.1 implies the existence of the integral (4.4.2).

**Theorem 4.4.1** If  $\tilde{P}$  has the distribution function  $F(P|\psi) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under consecutive sampling, then the expected gain,  $\bar{g}(\psi)$ , is given by

$$\bar{g}(\psi) = \sum_{i=1}^N \sum_{j=1}^N \bar{\pi}_i(T_{ij}(\psi)) r_{ij}(\psi) \quad i,j \in \Psi \quad (4.4.3)$$

where  $\bar{\pi}_i(\psi)$  is defined by equation (4.2.33).

**Proof.** By equation (4.4.1),

$$\bar{g}(\psi) = \sum_{i=1}^N \sum_{j=1}^N r_{ij} \int_{\mathcal{S}_N} \pi_i(P) p_{ij} dF(P|\psi). \quad (4.4.4)$$

Application of Lemma 2.3.2 yields (4.4.3). Q.E.D.

**Theorem 4.4.2** If  $\tilde{P}$  has the distribution function  $F(P|\psi) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under consecutive sampling, then the variance of  $g(\tilde{P})$  is

$$\text{var}[g(\tilde{P})|\psi] =$$

$$\sum_{i,j,k,m=1}^N \bar{p}_{ij}(\psi) r_{ij} r_{km} [\bar{\pi}_{km}(T_{ij}(\psi)) \bar{\pi}_{ik}(T_{km}(T_{ij}(\psi))) - \bar{p}_{km}(\psi) \bar{\pi}_i(T_{ij}(\psi)) \bar{\pi}_k(T_{km}(\psi))] \quad \psi \in \Psi \quad (4.4.5)$$

where  $\bar{\pi}_{ij}(\psi)$  is defined by equation (4.2.52).



**Proof.** The mean square of  $\underline{g}(\tilde{P})$  is, using Lemma 2.3.2 and equation (4.2.3)

$$\begin{aligned}
 E[\underline{g}^2(\tilde{P})|\psi] &= \sum_{i,j,k,m=1}^N r_{ij} r_{km} \int_{\mathcal{S}_N} \pi_i \pi_k p_{ij} p_{km} dP(P|\psi) \\
 &= \sum_{i,j,k,m=1}^N r_{ij} r_{km} \bar{p}_{ij}(\psi) \bar{p}_{km}(T_{ij}(\psi)) \int_{\mathcal{S}_N} \pi_i \pi_k dP(P|\pi_{km}(T_{ij}(\psi))) \\
 &= \sum_{i,j,k,m=1}^N r_{ij} r_{km} \bar{p}_{ij}(\psi) \bar{p}_{km}(T_{ij}(\psi)) \bar{\pi}_{ik}(T_{km}(T_{ij}(\psi))). 
 \end{aligned} \tag{4.4.6}$$

The existence of  $E[\underline{g}^2(\tilde{P})|\psi]$  follows from Lemma 4.2.1 and the continuity of the bounded function  $\pi_i(P) \pi_k(P) p_{ij} p_{km}$  on  $\mathcal{S}_N$ . From (4.4.3), the square mean of  $\underline{g}(\tilde{P})$  is

$$[\underline{g}(\psi)]^2 = \sum_{i,j,k,m=1}^N r_{ij} r_{km} \bar{p}_{ij}(\psi) \bar{p}_{km}(\psi) \bar{\pi}_i(T_{ij}(\psi)) \bar{\pi}_k(T_{km}(\psi)). \tag{4.4.7}$$

Subtracting (4.4.7) from (4.4.6), equation (4.4.5) is obtained. Q.E.D.

**Theorem 4.4.3** If  $\tilde{P}$  has the distribution function  $F(P|\psi) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is continuous in  $\psi$ , then the expectations  $\bar{g}(\psi)$  and  $\text{var}[\underline{g}(\tilde{P})|\psi]$  are continuous functions of  $\psi$  on  $\Psi$ .

**Proof.** The theorem follows immediately from Theorem 2.4.3. Q.E.D.

#### 4.4.2 Approximations to the Mean and Covariance Matrix of $\underline{V}(\tilde{P})$ .

The preceding results can be used to approximate the mean and the covariance matrix of the discounted reward vector  $\underline{V}(\tilde{P})$  discussed in Section 4.3. We assume throughout that the prior distribution function



of  $\tilde{P}$  is  $F(P|t) \in \mathcal{F}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under consecutive sampling.

The expected value of  $v_1(\tilde{P})$  is  $\tilde{v}_1(t)$ , which is given by equation (4.3.7) in terms of the mean n-step transition probabilities,  $\tilde{p}_{ij}^{(n)}(t)$ .

Since  $\lim_{n \rightarrow \infty} \tilde{p}_{ij}^{(n)}(t) = \bar{\pi}_j(t)$ , we can replace  $\tilde{p}_{ij}^{(n)}(t)$  by  $\bar{\pi}_j(t)$  in (4.3.7) for all  $n$  larger than some integer  $n^*$  to obtain the approximation

$$\tilde{v}_1(t) \doteq$$

$$\sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N \tilde{p}_{ij}^{(n)}(T_{jk}(t)) \tilde{p}_{jk}(t) r_{jk} + \frac{\beta^{n^*+1}}{1-\beta} \sum_{j=1}^N \sum_{k=1}^N \bar{\pi}_j(T_{jk}(t)) \tilde{p}_{jk}(t) r_{jk}, \quad (4.4.8)$$

or, using (4.4.3)

$$\tilde{v}_1(t) \doteq \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N \tilde{p}_{ij}^{(n)}(T_{jk}(t)) \tilde{p}_{jk}(t) r_{jk} + \frac{\beta^{n^*+1}}{1-\beta} \tilde{g}(t). \quad (4.4.9)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $t \in \mathbb{R}$   
 $0 \leq \beta < 1$

The error incurred when the approximation (4.4.9) is used is

$$e_1(n^*, t) = \sum_{n=n^*+1}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N [\tilde{p}_{ij}^{(n)}(T_{jk}(t)) - \bar{\pi}_j(T_{jk}(t))] \tilde{p}_{jk}(t) r_{jk}. \quad (4.4.10)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $t \in \mathbb{R}$   
 $0 \leq \beta < 1$

Since

$$|\tilde{p}_{ij}^{(n)}(T_{jk}(t)) - \bar{\pi}_j(T_{jk}(t))| \leq 1, \quad (4.4.11)$$

$i, j=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $t \in \mathbb{R}$

$e_1(n^*, t)$  can be bounded by



$$|a_1(n^*, \psi)| \leq \frac{\beta^{n^*} + 1}{1-\beta} \sum_{j=1}^N |\bar{q}_j(\psi)|, \quad (4.4.12)$$

$\begin{matrix} i=1, \dots, N \\ n^*=0, 1, 2, \dots \\ \psi \in \Psi \\ 0 \leq \beta < 1 \end{matrix}$

where

$$\bar{q}_j(\psi) = \sum_{k=1}^N \bar{p}_{jk}(\psi) r_{jk} \quad \begin{matrix} j=1, \dots, N \\ \psi \in \Psi \end{matrix}$$

is the mean one-step transition reward.

The bound (4.4.12) is conservative; tighter bounds require a tighter bound on  $|\bar{p}_{ij}^{(n)}(\psi) - \bar{\pi}_j(\psi)|$  than that provided by (4.4.11). This is a problem for future investigation.

The covariance between  $V_1(\tilde{P})$  and  $V_j(\tilde{P})$  is given by equation (4.3.22).

By Theorem 4.2.8,

$$\lim_{\substack{n \rightarrow \infty \\ v \rightarrow \infty}} E[\bar{p}_{ia}^{(n)} \bar{p}_{jm}^{(v)} | \psi] = \bar{\pi}_{am}(\psi). \quad \begin{matrix} i, j, a, m = 1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (4.4.13)$$

For all  $n > n^*$  and  $v > v^*$ , let us use the approximations

$$E[\bar{p}_{ia}^{(n)} \bar{p}_{jm}^{(v)} | T_{ay}(T_{ak}(\psi))] \approx \bar{\pi}_{am}(T_{ay}(T_{ak}(\psi))) \quad (4.4.14)$$

$$\bar{p}_{ia}^{(n)}(T_{ak}(\psi)) \bar{p}_{jm}^{(v)}(T_{ay}(\psi)) \approx \bar{\pi}_a(T_{ak}(\psi)) \bar{\pi}_m(T_{ay}(\psi)) \quad (4.4.15)$$

in equation (4.3.22). Using (4.4.5), we then have

$$\begin{aligned} \text{cov}[V_1(\tilde{P}), V_j(\tilde{P}) | \psi] &\approx \sum_{n=0}^{n^*} \sum_{v=0}^{v^*} \beta^{n+v} \sum_{a,y,k,m=1}^N \bar{p}_{ak}(\psi) r_{ak} r_{ay} \\ &\times \left[ E[\bar{p}_{ia}^{(n)} \bar{p}_{jm}^{(v)} | T_{ay}(T_{ak}(\psi))] \bar{p}_{ay}^{(n)}(T_{ak}(\psi)) - \bar{p}_{ia}^{(n)}(T_{ak}(\psi)) \bar{p}_{jm}^{(v)}(T_{ay}(\psi)) \bar{p}_{ay}^{(n)}(\psi) \right. \\ &\left. + \frac{\beta^{n^*+v^*+2}}{(1-\beta)^2} \text{var}[g(\tilde{P}) | \psi] \right]. \quad \begin{matrix} i, j = 1, \dots, N \\ \psi \in \Psi \end{matrix} \quad (4.4.16) \end{aligned}$$

The error involved in using (4.4.16) to approximate  $\text{cov}[V_1(\tilde{P}), V_j(\tilde{P}) | \psi]$



$$e_{ij}(n^*, v^*; \gamma) = \sum_{m=n+1}^{\infty} \sum_{v=v+1}^{\infty} \beta^{2m+v} \sum_{a,y,k,u=1}^N \tilde{p}_{ak}(\gamma) r_{ak} r_{ay}$$

$$\begin{aligned} & \times \left[ [\mathbb{E} [\tilde{p}_{ia}^{(n)} \tilde{p}_{ja}^{(v)} | T_{ay}(T_{ak}(\gamma))] - \bar{\pi}_a(T_{ay}(T_{ak}(\gamma))) \tilde{p}_{ay}(T_{ak}(\gamma)) \right. \\ & \left. - [\tilde{p}_{ia}^{(n)}(T_{ak}(\gamma)) \tilde{p}_{ja}^{(v)}(T_{ay}(\gamma)) - \bar{\pi}_a(T_{ak}(\gamma)) \bar{\pi}_a(T_{ay}(\gamma))] \tilde{p}_{ay}(\gamma) \right]. \end{aligned} \quad (4.4.17)$$

A conservative bound on  $e_{ij}(n^*, v^*; \gamma)$  is

$$|e_{ij}(n^*, v^*; \gamma)| \leq \frac{\beta^{n^*+v^*+2}}{(1-\beta)^2} \sum_{a,y,k,u=1}^N \tilde{p}_{ak}(\gamma) |r_{ak}| \cdot |r_{ay}| [\tilde{p}_{ay}(T_{ak}(\gamma)) + \tilde{p}_{ay}(\gamma)] \quad (4.4.18)$$

**4.4.3 An Abelian Theorem.** We conclude this chapter with a theorem which relates  $\tilde{g}(\gamma)$  to  $\tilde{V}_1(\gamma)$ . Some results from the theory of summation of divergent series are required and are summarized here without proof.\*

Let  $\{a_n\} = \{a_0, a_1, a_2, \dots\}$  be a sequence of real numbers. Let

$$t_n = \frac{1}{n+1} \sum_{v=0}^n a_v. \quad n=0, 1, 2, \dots \quad (4.4.19)$$

If  $\lim_{n \rightarrow \infty} t_n$  exists and is equal to  $t$ , then the sequence  $\{a_n\}$  is said to be Cesaro-summable, or  $\mathcal{L}_1$ -summable, to  $t$ . If the  $\mathcal{C}_1$ -sum of  $\{a_n\}$  exists and is equal to  $t$ , then a theorem due to Abel states that  $\lim_{\beta \downarrow 1^-} (1-\beta) \sum_{v=0}^{\infty} a_v \beta^v$  exists and is equal to  $t$ .

**Theorem 4.4.4** Let  $\tilde{P}$  have the distribution function  $F(P|\gamma) \in \mathcal{D}$ , a family of distributions continuous on the boundary of  $\mathcal{S}_N$  which is closed under the consecutive sampling rule. Let  $\tilde{V}_2(\gamma) = \tilde{V}_1(\beta, \gamma)$  and  $\tilde{g}(\gamma)$  be

\* See, for example, Knopp [27].



defined by equations (4.3.2) and (4.4.2), respectively. Then

$$\lim_{\beta \rightarrow 1^-} (1-\beta) \tilde{V}_1(\beta, \gamma) = \tilde{g}(\gamma). \quad i=1, \dots, N \quad (4.4.20)$$

**Proof.** We first show that the  $C_1$ -sum of the sequence  $\left\{ \tilde{p}_{1j}^{(n)}(\gamma) \right\}$  exists and is equal to  $\bar{\pi}_j(\gamma)$  for  $j=1, \dots, N$  and any  $\gamma \in \Gamma$ . Let

$$t_{1j}(n, \gamma) = \frac{1}{n+1} \sum_{v=0}^n \tilde{p}_{1j}^{(v)}(\gamma). \quad i, j=1, \dots, N \quad (4.4.21)$$

$n=0, 1, 2, \dots$   
 $\gamma \in \Gamma$

Let  $\epsilon > 0$  be given. Choose  $n^*$  such that, for fixed indices  $i$  and  $j$  and fixed  $\gamma \in \Gamma$ ,

$$\left| \tilde{p}_{1j}^{(n)}(\gamma) - \bar{\pi}_j(\gamma) \right| < \frac{\epsilon}{2}. \quad n > n^* \quad (4.4.22)$$

Then, for  $n > n^*$ ,

$$\begin{aligned} |t_{1j}(n, \gamma) - \bar{\pi}_j(\gamma)| &\leq \frac{1}{n+1} \sum_{v=0}^n |\tilde{p}_{1j}^{(v)}(\gamma) - \bar{\pi}_j(\gamma)| \\ &< \frac{1}{n+1} \sum_{v=0}^{n^*} |\tilde{p}_{1j}^{(v)}(\gamma) - \bar{\pi}_j(\gamma)| + \frac{\epsilon}{2} \end{aligned} \quad (4.4.23)$$

By choosing an integer  $\mu > n^*$  such that, if  $n > \mu$ ,

$$0 \leq \frac{1}{n+1} \sum_{v=0}^n \left| \tilde{p}_{1j}^{(v)}(\gamma) - \bar{\pi}_j(\gamma) \right| < \frac{\epsilon}{2}, \quad (4.4.24)$$

we have, for  $n$ ,

$$\left| t_{1j}(n, \gamma) - \bar{\pi}_j(\gamma) \right| < \epsilon \quad (4.4.25)$$

and, therefore,  $\lim_{n \rightarrow \infty} t_{1j}(n, \gamma) = \bar{\pi}_j(\gamma)$ , proving the assertion.

Using equation (4.3.7),

$$\lim_{\beta \rightarrow 1^-} (1-\beta) \tilde{V}_1(\beta, \gamma) = \sum_{j=1}^N \sum_{k=1}^N \tilde{p}_{jk}(\gamma) r_{jk} \lim_{\beta \rightarrow 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n \tilde{p}_{1j}^{(n)}(T_{jk}(\gamma)) \quad (4.4.26)$$



$\sigma_1$ -summability of  $\{\tilde{p}_{ij}^{(n)}(T_{jk}(\gamma))\}$  to  $\bar{\pi}_j(T_{jk}(\gamma))$  implies that

$\lim_{\beta \rightarrow 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n \tilde{p}_{ij}^{(n)}(T_{jk}(\gamma))$  exists and is equal to  $\bar{\pi}_j(T_{jk}(\gamma))$ .

Thus, using equation (4.4.3),

$$\begin{aligned} \lim_{\beta \rightarrow 1^-} (1-\beta) \tilde{v}_i(\beta, \gamma) &= \sum_{j=1}^N \sum_{k=1}^N \bar{\pi}_j(T_{jk}(\gamma)) \tilde{p}_{jk}(\gamma) r_{jk} \\ &= \tilde{g}(\gamma). \quad \begin{matrix} i=1, \dots, N \\ \gamma \in \Gamma \end{matrix} \quad (4.4.27) \end{aligned}$$

Q.E.D.



## CHAPTER 5

### TERMINAL CONTROL PROBLEMS

In this chapter we consider sequential sampling models of a Markov chain with alternatives in which there is an explicit sampling cost. This leads to a distinction between sampling the process and using the process. When the process is sampling, the sequence of states occupied by the Markov chain during the sampling period is made known to the decision-maker, who then uses this information to update his prior distribution on  $\underline{\hat{P}}$ . During the sampling period the process earns transition rewards as specified by  $\underline{Q}$  and sampling costs are incurred.

On the other hand, if the process is used over a period of  $n$  transitions, it earns transition rewards, but the decision-maker is permitted to know only the initial state and the final state of the sample sequence. The only sample cost incurred is that for observing the state of the system after the  $n$ th transition.

It is reasonable to expect that, after a finite amount of sampling, the prior distribution of  $\underline{\hat{P}}$  will be sufficiently tight that the best course of action for the decision-maker will be to cease sampling and operate the process under some fixed terminal policy indefinitely. Hence, the models of this chapter are called terminal control models. In the following sections we show that such a terminal decision point occurs with probability one in an optimal sampling strategy.

Terminal control models are applicable, in general, to any Markov chain with alternatives in which rewards are earned independently of the decision-



maker's knowledge of the sequence of states occupied by the system and in which it is possible for him to determine the state of the system at any time, for a non-zero cost. A specific example of such a process is a Markov chain model of consumer brand-switching behavior, where a survey must be made to determine the current state of the market.

Two-action sequential sampling problems with independent identically distributed observations have been examined from the Bayesian point of view by Wetherill [40]. A similar problem with Markov-dependent observations was recently considered by Ehat [9].

In Section 5.1 we examine Model I, a discounted terminal control model in which the decision-maker must sample at every transition of the process until a terminal decision point is reached. This model is formulated as a set of functional equations and it is shown that a terminal decision point is reached with probability one in an optimal sampling strategy. It is shown, in Section 5.2, that there exists a unique solution to these equations and a method of successive approximations is introduced. This model is generalized in Section 5.3, where Model II is introduced. Model II is a discounted terminal control model in which the decision-maker can either sample or use the process until a terminal decision is made. Approximate methods of making terminal decisions are discussed in Section 5.4. Models of undiscounted processes are introduced in Section 5.5 and the chapter concludes with a brief discussion of set-up costs.

### 5.1 Discounted Processes. Model I.

Consider a Markov chain with alternatives which has the reward matrix  $\underline{R} = [r_{ij}^k]$ . At each transition the decision-maker can either sample the process or choose a terminal policy under which the system is to be



operated indefinitely. Let  $c_i > 0$  be the cost of observing the system and finding it in state  $i$  ( $i=1, \dots, N$ ).<sup>\*</sup> The cost of any sampling strategy is, therefore, a random variable before it is executed. Assuming that the interval between transitions is constant, we may use this interval as the unit of time. Let  $\beta$  be the present value of a unit reward received one unit of time in the future ( $0 \leq \beta < 1$ ). We shall seek a sampling strategy which maximizes the expected total discounted reward over an infinite period.

When the decision-maker chooses to sample we clearly have a case of consecutive sampling. Thus, it is assumed that the prior distribution function of  $\tilde{P}$  is  $H(\underline{\theta}|\gamma) \in \mathcal{H}$ , a family of distributions closed under consecutive sampling. Let  $(i, \gamma)$  denote the generalized state of the system ( $i=1, \dots, N$ ;  $\gamma \in \Gamma$ ) and let  $v_i(\gamma)$  be the supremum of the expected total discounted reward over an infinite period if the system starts from the generalized state  $(i, \gamma)$ . In Theorem 5.1.1 it will be shown that an optimal sampling strategy exists and, therefore, that  $v_i(\gamma)$  is the maximum expected discounted reward over an infinite period.

If, when in state  $(i, \gamma)$ , it is decided to sample the  $k$ th alternative and the system then makes a transition to state  $j$ , the supremum of the posterior expected reward is

$$r_{ij}^k = \beta c_j + \beta v_j(T_{ij}^k(\gamma)). \quad (5.1.1)$$

\* Model I is easily generalized to allow the sampling cost to be  $c_{ij}^k$ , the cost of observing a transition from state  $i$  to state  $j$  under the  $k$ th alternative in state  $i$ . In this case equation (5.1.3) below is

$$\tilde{c}_i^k(\gamma) = \sum_{j=1}^N P_{ij}(\gamma) c_{ij}^k.$$

Model II, however, requires that the sampling cost be independent of the state from which the transition originated.



The probability of the sample outcome  $j$ , unconditional with respect to the prior distribution, given that the system is in state  $(i, \tau)$  and alternative  $k$  is in use, is the prior expected value of  $\hat{P}_{ij}^k$ .

$$\hat{P}_{ij}^k(\tau) = \int_{\Omega_{N,N}} P_{ij}^k dF(P|\tau). \quad (5.1.2)$$

Let  $\hat{q}_i^k(\tau)$  be the mean one-step transition reward defined by equation (3.1.4) and let

$$\hat{c}_i^k(\tau) = \sum_{j=1}^N \hat{P}_{ij}^k(\tau) c_j \quad (5.1.3)$$

be the expected cost of sampling alternative  $k$  when in the state  $(i, \tau)$ . Then, if it is decided to sample the process on the next transition, the supremum of the prior expected reward is

$$\max_{1 \leq k \leq K_2} \left\{ \hat{q}_i^k(\tau) + \beta \hat{c}_i^k(\tau) + \beta \sum_{j=1}^N \hat{P}_{ij}^k(\tau) v_j(\hat{r}_{ij}^k(\tau)) \right\} \underset{\begin{array}{l} i=1, \dots, N \\ \tau \in \mathbb{T} \\ \Sigma \in \Sigma \end{array}}{\text{max}} \quad (5.1.4)$$

Suppose, on the other hand, it is decided to cease sampling and to operate the process indefinitely under the policy  $\Sigma$ . Let

$$\hat{V}_i(\Sigma, \tau) = \int_{\Omega_N} V_i(P) dF_{\Sigma}(P|\tau) \quad \underset{\begin{array}{l} i=1, \dots, N \\ \tau \in \mathbb{T} \\ \Sigma \in \Sigma \end{array}}{\text{def}} \quad (5.1.5)$$

be the unconditional expected discounted reward over an infinite period<sup>\*</sup> when the policy  $\Sigma$  is used and the system starts from  $(i, \tau)$ . Then, if it is decided to cease sampling, the maximum prior expected reward is

$$\max_{\Sigma \in \Sigma} \left\{ \hat{V}_i(\Sigma, \tau) \right\}. \quad \underset{\begin{array}{l} i=1, \dots, N \\ \tau \in \mathbb{T} \end{array}}{\text{max}} \quad (5.1.6)$$

The maximum exists in (5.1.6) since  $\Sigma$  is a finite set.

Using equations (5.1.4) and (5.1.6), we have the following set of

\* Cf. Section 4.3.



functional equations which must be satisfied by the vector function,

$$v(\tau) = (v_1(\tau), \dots, v_N(\tau))$$

$$v_1(\tau) = \max_{\substack{i \in k \in \mathcal{A}_1 \\ \sigma \in \Sigma}} \left\{ \bar{q}_1^k(\tau) - \beta \bar{v}_1^k(\tau) + \beta \sum_{j=1}^N \bar{p}_{1,j}^k(\tau) v_j(T_{1,j}^k(\tau)) \right\}$$

$$\max_{\substack{i=1, \dots, N \\ \tau \in \Gamma \\ 0 \leq \beta < 1}} \left\{ \bar{v}_1(\sigma, \tau) \right\}$$

(5.1.7)

It is to be noted that the same symbol,  $v_1(\tau)$ , is used in equation (3.1.5) and equation (5.1.7) to represent two distinct functions. This has been done to simplify to some extent a necessarily complicated notation. The meaning of the symbol  $v_1(\tau)$  will always be clear from its context.

We now show that an optimal sampling strategy exists, making use of the definitions and notation of Section 3.1.

**Theorem 5.1.1** Let  $v_1(\tau, d)$  be the expected total discounted reward in Model I when the system starts in the generalised state  $(i, \tau)$  and the sampling strategy  $d \in \mathcal{D}_1$  is used. Let

$$v_1(\tau) = \sup_{d \in \mathcal{D}_1} \left\{ v_1(\tau, d) \right\} \quad i=1, \dots, N \quad (5.1.8)$$

Then there is a sampling strategy  $d^* \in \mathcal{D}_1$  such that

$$v_1(\tau) = v_1(\tau, d^*). \quad i=1, \dots, N \quad (5.1.9)$$

**Proof.** Consider the adaptive control problem of Section 3.1. Let  $\tilde{\mathcal{D}}_1$  be the set of all possible sampling strategies,  $\tilde{d}$ , in the adaptive control problem when the system starts from state  $i$ . In Model I, if

\* See esp., pp. 42, 43.



$d \in D_1$ , it is clear that  $d$  is a possible strategy in the adaptive control problem; hence,  $D_1 \subset \tilde{D}_1$  ( $i=1, \dots, N$ ). Suppose  $d \in \tilde{D}_1$ . Then  $\tilde{d}$  either prescribes a fixed policy,  $\underline{\Sigma}_i$ , for use on every transition  $n > n^*$ , for some integer  $n^*$ , or else not. In the first case,  $\tilde{d}$  is clearly a possible strategy for Model I. In the second case,  $\tilde{d}$  is also a possible strategy for Model I, viz., a strategy in which a terminal decision point is never reached under some or all possible sample histories. Thus,  $\tilde{D}_1 \subset D_1$  and, therefore,  $\tilde{D}_1 = D_1$ . The proof of Theorem 3.1.1 is valid for an arbitrary reward structure, provided that the reward per transition is bounded; thus, the remainder of the proof of Theorem 5.1.1 follows the proof of Theorem 3.1.1 and will not be duplicated here. Q.E.D.

We now show that, with probability one, a terminal decision point is reached in Model I if an optimal sampling strategy is used. Let  $\underline{Q}$  denote the true state of nature;  $\underline{Q}$  is assumed to be positive, as defined by equation (2.3.26).

**Theorem 5.1.2** In Model I, if the true state of nature,  $\underline{Q}$ , is a positive matrix and if the terminal functions  $\tilde{V}_i(\underline{\Sigma}, \cdot)$  are continuous in  $\cdot$  ( $i=1, \dots, N$ ;  $\underline{\Sigma} \in \Sigma$ ), then, with probability one, a terminal decision point is reached in an optimal sampling strategy.

**Proof.** The proof is by contradiction. Assume there is an optimal sampling strategy in which a terminal decision is never made. Then the process is sampled infinitely often under the consecutive sampling rule and at least one state,  $i$ , is entered infinitely often. Since at least one alternative,  $k$ , must be used infinitely often when in state  $i$ , Lemma 2.3.7 and the positivity of  $\underline{Q}$  imply that every state is entered infinitely often, and, therefore, that at least one alternative in each state is



sampled infinitely often. Since the system is operating under an optimal strategy any alternative which is sampled a finite number of times is dominated by other alternatives after a finite number of transitions and can be eliminated from further consideration. Thus, we may assume, without loss of generality, that all alternatives are sampled infinitely often. By Theorem 2.3.8, the mass of the posterior distribution of  $\underline{\Phi}$  tends, with probability one, to concentrate at  $\underline{Q}$  as  $n$ , the number of trials, goes to infinity. That is, for any  $\epsilon > 0$ , if  $P_n$  is defined by equation (2.3.34),

$$\lim_{n \rightarrow \infty} P_n [ |\underline{\Phi} - \underline{Q}| < \frac{\epsilon}{2} ] = 1, \quad (5.1.10)$$

the limit holding with probability one. Let  $H(\underline{\Phi} | \tau^*)$  be the distribution function which places the unit mass of probability on  $\underline{Q}$ . Then, with probability one,  $\tau \rightarrow \tau^*$  as  $n \rightarrow \infty$  and equation (5.1.4) becomes<sup>4</sup>

$$v_i(\tau^*) = \max_{1 \leq k \leq K_1} \left\{ \sum_{j=1}^N p_{ij}^{(k)} (\tau^*) [r_{ij}^k - \beta c_j + \beta v_j(\tau^*)] \right\}, \quad (5.1.11)$$

$i=1, \dots, N$

Equation (5.1.11) is a sequential decision problem in which the decision-maker is certain about the transition probabilities and has been studied by several authors. Blackwell [11] has shown that an optimal strategy exists for (5.1.11) in which a fixed policy,  $\Sigma \in \Sigma$ , is used at every transition. Howard [22] has shown that the expected reward under this strategy is, in the notation of this proof,

$$v_i(\tau^*) = \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N q^{(n)}(j) \sum_{j'k'} q_{jj'}^{(n)} \frac{r_{jj'}^k}{\beta^k} [\tau_{jk'}^k - \beta c_k], \quad (5.1.12)$$

$i=1, \dots, N$

where  $q_{ij}^{(n)}(\Sigma)$  is the  $(i,j)$ th element of  $[q(\Sigma)]^n$ . But, as  $n \rightarrow \infty$ ,

<sup>4</sup> The argument is valid even if the distribution which places the unit mass of probability on  $\underline{Q}$  is not a member of  $\Psi$ .



$\tilde{v}_i(\underline{\sigma}, \dot{\tau}) \rightarrow \tilde{v}_i(\underline{\sigma}, \tau^*)$  with probability one. Let  $\tilde{v}_i(\underline{\sigma}^*, \tau^*) = \max_{\underline{\sigma} \in \Sigma} \{\tilde{v}_i(\underline{\sigma}, \tau^*)\}$ . Since  $c_j > 0$  ( $j=1, \dots, N$ ) equation (4.3.1) implies the contradiction

$$\begin{aligned} v_i(\tau^*) &< \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N q_{ij}^{(n)}(\underline{\sigma}) c_j \frac{\sigma_j}{jk} r \frac{\sigma_j}{jk} \\ &= \tilde{v}_i(\underline{\sigma}, \tau^*) \\ &\leq \tilde{v}_i(\underline{\sigma}^*, \tau^*). \quad i=1, \dots, N \quad (5.1.13) \end{aligned}$$

Therefore, with probability one, a terminal decision point is reached after a finite number of transitions. Q.E.D.

## 5.2 Existence and Uniqueness of Solutions. Successive Approximations.

Let the sequence of vector functions,  $\mathbf{v}(n, \dot{\tau})$ , where

$\mathbf{v}(n, \dot{\tau}) = (v_1(n, \dot{\tau}), \dots, v_N(n, \dot{\tau}))$ , be defined by the equations

$$v_i(n+1, \dot{\tau}) = \max \left[ \begin{array}{l} \max_{1 \leq k \leq K_1} \left\{ \bar{c}_i^k(\dot{\tau}) - \beta \bar{c}_i^k(\dot{\tau}) + \beta \sum_{j=1}^N \bar{p}_{ij}^k(\dot{\tau}) v_j(n, \tau_{2,j}^k(\dot{\tau})) \right\} \\ \max_{\underline{\sigma} \in \Sigma} \left\{ \tilde{v}_i(\underline{\sigma}, \dot{\tau}) \right\} \end{array} \right] \quad \begin{array}{l} i=1, \dots, N \quad (5.2.1a) \\ n=1, 2, \dots \\ \dot{\tau} \in \mathbb{P} \\ 0 \leq \beta < 1 \end{array}$$

$$v_i(0, \dot{\tau}) = \max_{\underline{\sigma} \in \Sigma} \left\{ \tilde{v}_i(\underline{\sigma}, \dot{\tau}) \right\}. \quad i=1, \dots, N \quad (5.2.1b) \quad \dot{\tau} \in \mathbb{P}$$

Using equation (5.2.1), it is shown in this section that there exists a unique bounded solution to equation (5.1.7). Equation (5.2.1) can then be used as a computational tool to approximate this unique solution and, with this application in mind, a bound on the error,  $e_i(n, \dot{\tau}) = v_i(\dot{\tau}) - v_i(n, \dot{\tau})$ , is derived. With the aid of this bound, we show that



$\{y(n, \gamma)\} \rightarrow y(\gamma)$  uniformly in  $\gamma$ .

Lemma 5.2.1 If  $R^* = \max_{i,j,k} \{|r_{ij}^k|\}$  and  $C = \max_j \{c_j\}$ , then the functions  $v_i(n, \gamma)$  defined by (5.2.1) have the bound

$$|v_i(n, \gamma)| \leq \frac{R^* + \beta C}{1-\beta}. \quad \begin{array}{l} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{array} \quad (5.2.2)$$

Proof. By equations (4.3.5) and (4.3.6),

$$|\tilde{v}_i(\underline{\sigma}, \gamma)| \leq \frac{R^*}{1-\beta}, \quad \begin{array}{l} i=1, \dots, N \\ \gamma \in \Gamma \\ \underline{\sigma} \in \Sigma \end{array} \quad (5.2.3)$$

and, since  $C > 0$ ,

$$|v_i(0, \gamma)| < \frac{R^* + \beta C}{1-\beta}. \quad \begin{array}{l} i=1, \dots, N \\ \gamma \in \Gamma \end{array} \quad (5.2.4)$$

Assume that (5.2.2) holds for  $n$ . Then

$$\begin{aligned} |v_i(n+1, \gamma)| &\leq \max [R^* + \beta C + \frac{\beta R^* + \beta^2 C}{1-\beta}, \frac{R^*}{1-\beta}] \\ &= \frac{R^* + \beta C}{1-\beta}. \end{aligned} \quad (5.2.5)$$

Q.E.D.

Theorem 5.2.2 If  $y(n, \gamma)$  is defined by equation (5.2.1), then

$\{v_i(n, \gamma)\}$  is a monotone increasing sequence ( $i=1, \dots, N$ ;  $\gamma \in \Gamma$ ) and  $\lim_{n \rightarrow \infty} y(n, \gamma)$  exists and is a solution to (5.1.7).

Proof. We show inductively that  $\{v_i(n, \gamma)\}$  is a monotone increasing sequence. Since  $v_i(0, \gamma) = \max_{\underline{\sigma} \in \Sigma} \{\tilde{v}_i(\underline{\sigma}, \gamma)\}$  we have  $v_i(1, \gamma) \geq v_i(0, \gamma)$ . Assume that  $v_i(n, \gamma) \geq v_i(n-1, \gamma)$  for  $i=1, \dots, N$  and  $\gamma \in \Gamma$ . If  $v_i(n, \gamma) = \max_{\underline{\sigma} \in \Sigma} \{\tilde{v}_i(\underline{\sigma}, \gamma)\}$ , then  $v_i(n+1, \gamma) \geq v_i(n, \gamma)$ . Suppose that, for some  $k \in \{1, \dots, K_i\}$ ,



$$v_1(n, \gamma) = \bar{q}_1^k(\gamma) - \beta \bar{c}_1^k(\gamma) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) v_j(n-1, T_{1j}^k(\gamma)). \quad (5.2.6)$$

Then, since

$$v_1(n+1, \gamma) \geq \bar{q}_1^k(\gamma) - \beta \bar{c}_1^k(\gamma) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) v_j(n, T_{1j}^k(\gamma)), \quad (5.2.7)$$

we have

$$\begin{aligned} v_1(n+1, \gamma) - v_1(n, \gamma) &\geq \beta \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) [v_j(n, T_{1j}^k(\gamma)) - v_j(n-1, T_{1j}^k(\gamma))] \\ &\geq 0, \end{aligned} \quad (5.2.8)$$

proving the induction. By Lemma 5.2.1, the sequence  $\{v_1(n, \gamma)\}$  is bounded, hence,  $\lim_{n \rightarrow \infty} v_1(n, \gamma)$  exists ( $i=1, \dots, N$ ). That the limit is a solution of (5.1.7) is seen by letting  $n \rightarrow \infty$  in (5.2.1a). Q.E.D.

**Theorem 5.2.3** There is a unique bounded solution to equation (5.1.7).

**Proof.** It was shown in Theorem 5.2.3, that there is at least one bounded solution,  $x(\gamma)$ , to (5.1.7). Assume that  $y(\gamma) = (y_1(\gamma), \dots, y_N(\gamma))$  is also a bounded solution. Let

$$s_1^k(v, \infty, \gamma) = \bar{q}_1^k(\gamma) - \beta \bar{c}_1^k(\gamma) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) v_j(T_{1j}^k(\gamma)). \quad (5.2.9)$$

$k=1, \dots, K_1$   
 $i=1, \dots, N_1$   
 $\gamma \in \mathbb{R}$

Assume  $(i, \gamma)$  is fixed. There are four cases.

Case 1.  $v_1(\gamma) = \max_{\underline{\sigma} \in \Sigma} \{\bar{v}_1(\underline{\sigma}, \gamma)\}$  and  $w_1(\gamma) = \max_{\underline{\sigma} \in \Sigma} \{\bar{v}_1(\underline{\sigma}, \gamma)\}$ . Then  $v_1(\gamma) = w_1(\gamma)$ .

Case 2. For some  $a \in \{1, \dots, K_1\}$ ,  $v_1(\gamma) = s_1^a(v, \infty, \gamma)$  and  $w_1(\gamma) = \max_{\underline{\sigma} \in \Sigma} \{\bar{v}_1(\underline{\sigma}, \gamma)\}$ . Then

$$0 \leq v_1(\gamma) - w_1(\gamma) \leq s_1^a(v, \infty, \gamma) - s_1^a(v, \infty, \gamma). \quad (5.2.10)$$



Case 3. For some  $v \in \{1, \dots, K_1\}$ ,  $v_1(\gamma) = S_1^V(v, \infty, \gamma)$  and  $v_1(\gamma) = \max_{\sigma \in \Sigma} \{\tilde{v}_1(\sigma, \gamma)\}$ . Then

$$S_1^V(v, \infty, \gamma) = S_1^V(w, \infty, \gamma) \leq v_1(\gamma) = w_1(\gamma) \leq 0. \quad (5.2.11)$$

Case 4. For some indices  $a$  and  $b$  belonging to  $\{1, \dots, K_1\}$

$v_1(\gamma) = S_1^a(v, \infty, \gamma)$  and  $w_1(\gamma) = S_1^b(w, \infty, \gamma)$ . Then

$$S_1^b(v, \infty, \gamma) = S_1^b(w, \infty, \gamma) \leq v_1(\gamma) = w_1(\gamma) \leq S_1^a(v, \infty, \gamma) = S_1^a(w, \infty, \gamma). \quad (5.2.12)$$

Let  $k^*$  index the maximum of  $|S_1^a(v, \infty, \gamma) - S_1^a(w, \infty, \gamma)|$ ,

$|S_1^b(v, \infty, \gamma) - S_1^b(w, \infty, \gamma)|$ ,  $|S_1^a(v, \infty, \gamma) - S_1^a(w, \infty, \gamma)|$ , and

$|S_1^V(v, \infty, \gamma) - S_1^V(w, \infty, \gamma)|$ . Then, in all of the above cases,

$$\begin{aligned} |v_1(\gamma) - w_1(\gamma)| &\leq |S_1^{k^*}(v, \infty, \gamma) - S_1^{k^*}(w, \infty, \gamma)| \\ &\leq \beta \sum_{j=1}^N \tilde{p}_{1j}^{k^*}(\gamma) |v_j(T_{1j}^{k^*}(\gamma)) - w_j(T_{1j}^{k^*}(\gamma))|. \end{aligned} \quad (5.2.13)$$

Since  $v_1(\gamma)$  and  $w_1(\gamma)$  are both bounded, there exists a number,  $M > 0$ , such that

$$|v_1(\gamma) - w_1(\gamma)| < M.$$

Repeated application of (5.2.13) yields

$$|v_1(\gamma) - w_1(\gamma)| < \beta^n M. \quad \begin{matrix} n=0, 1, 2, \dots \\ \gamma \in \Gamma \end{matrix} \quad (5.2.14)$$

Since  $0 \leq \beta < 1$ , (5.2.14) implies  $v_1(\gamma) = w_1(\gamma)$ . Q.E.D.

We now derive a bound on the error of the  $n$ th approximant,  $v_1(n, \gamma)$ .

Let

$$S_1^k(v, n, \gamma) = \tilde{q}_1^k(\gamma) + \beta \sum_{j=1}^N \tilde{p}_{1j}^k(\gamma) v_j(n, T_{1j}^k(\gamma)). \quad (5.2.15)$$

$\begin{matrix} k=1, \dots, K_1 \\ j=1, \dots, N \\ n=0, 1, 2, \dots \\ \gamma \in \Gamma \end{matrix}$



**Theorem 5.2.4** Let the error of the  $n$ th approximant  $v_1(n, \gamma)$  be defined as

$$e_1(n, \gamma) = v_1(\gamma) - v_1(n, \gamma), \quad (5.2.16)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\gamma \in \mathbb{R}$

where  $v_1(n, \gamma)$  is defined by (5.2.1) and  $v_1(\gamma)$  is the unique bounded solution of (5.1.7). Let  $R$  and  $r$  be defined by equation (3.3.3). Then  $e_1(n, \gamma)$  has the bounds

$$0 \leq e_1(n, \gamma) \leq \beta^n \frac{R-r}{1-\beta}. \quad (5.2.17)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\gamma \in \mathbb{R}$   
 $0 \leq \beta < 1$

**Proof.** By Theorem 5.2.2,  $\{v_1(n, \gamma)\}$  is a monotone increasing sequence with the limit  $v_1(\gamma)$ ; hence,  $e_1(n, \gamma) \geq 0$  ( $n=0, 1, 2, \dots$ ). The remainder of the inequality (5.2.17) is proved by induction.

We first establish that  $v_1(\gamma) \leq \frac{R}{1-\beta}$ . Since  $\tilde{v}_1(\underline{\Sigma}, \gamma) \leq R \sum_{n=0}^{\infty} \beta^n = \frac{R}{1-\beta}$  for all  $\underline{\Sigma} \in \Sigma$ , we have

$$v_1(0, \gamma) \leq \frac{R}{1-\beta}. \quad (5.2.18)$$

Assume  $v_1(n, \gamma) \leq \frac{R}{1-\beta}$ . Then, since  $e_j > 0$  ( $j=1, \dots, N$ ),

$$v_1(n+1, \gamma) \leq \max [R + \frac{\beta R}{1-\beta}, \frac{R}{1-\beta}] = \frac{R}{1-\beta} \quad (5.2.19)$$

and, by induction,

$$v_1(n, \gamma) \leq \frac{R}{1-\beta}. \quad (5.2.20)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\gamma \in \mathbb{R}$

Thus,  $v_1(\gamma) = \lim_{n \rightarrow \infty} v_1(n, \gamma) \leq \frac{R}{1-\beta}$ , proving the assertion.

Since  $\tilde{v}_1(\underline{\Sigma}, \gamma) \geq \frac{R}{1-\beta}$  for all  $\underline{\Sigma} \in \Sigma$ , we have



$$e_1(0, \psi) \leq \frac{R\beta}{1-\beta} \cdot \begin{matrix} i=1, \dots, N \\ \psi \in \mathbb{P} \end{matrix} \quad (5.2.21)$$

and (5.2.17) holds for  $n = 0$ . Assume the equation is valid for  $n$ . Then, arguing as in the proof of Theorem 5.2.3, there is an index  $k \in \{1, \dots, R_1\}$  such that

$$\begin{aligned} e_1(n+1, \psi) &= |e_1(n+i, \psi)| \leq |S_1^k(v, \infty, \psi) - S_1^k(v, n, \psi)| \\ &\leq \beta \sum_{j=1}^N p_{1j}^k(\psi) |e_j(n, t_{1j}^k(\psi))| \\ &\leq \beta^{n+1} \frac{R\beta}{1-\beta}. \end{aligned} \quad (5.2.22)$$

Q.E.D.

**Corollary 5.2.5** If  $y(n, \psi)$  is defined by equation (5.2.1) and  $y(\psi)$  is the unique bounded solution of equation (5.1.7), then  $\{y(n, \psi)\} \rightarrow y(\psi)$  uniformly in  $\psi$ .

**Proof.** Since the error bound (5.2.17) is independent of  $\psi$ , the corollary follows from Theorem 5.2.4. Q.E.D.

**Theorem 5.2.6** If the prior distribution function of  $\tilde{\Psi}$  is  $H(\underline{P}|\psi) \in \mathcal{H}$ , a family of distributions continuous in  $\psi$  which is closed under consecutive sampling, then  $y(\psi)$ , the unique bounded solution of (5.1.7), is a continuous function of  $\psi$ .

**Proof.** Since  $\mathcal{H}$  is continuous in  $\psi$ ,  $v_1(0, \psi) = \max_{\sigma \in \Sigma} \{\tilde{V}_1(\sigma, \psi)\}$  is a continuous function of  $\psi$ . Moreover,  $p_{1j}^k(\psi)$  is continuous. Thus, by induction,  $v_1(n, \psi)$  is continuous for  $i=1, \dots, N$  and  $n=0, 1, 2, \dots$ . Since  $\{v_1(n, \psi)\} \rightarrow v_1(\psi)$  uniformly in  $\psi$ ,  $v_1(\psi)$  is continuous ( $i=1, \dots, N$ ). Q.E.D.



### 5.3 Measured Processes. Model II.

In considering Model I it is immediately apparent that the maximum expected reward will not be decreased--and may be increased--if we stop sampling in some states while continuing to sample in others. For example, if the marginal prior distribution of  $\tilde{p}_k^i$  is loose, while the marginal prior distribution of the remaining ( $K-1$ ) rows of  $\tilde{\underline{P}}$  is tight, it may be profitable to sample only when the system is in state  $i$ . Model II admits this additional option.

As in Section 5.1, let  $v_1(\gamma)$  be the supremum of the expected total discounted reward over an infinite period when the system starts from the generalised state  $(i, \gamma)$ . It is assumed that the decision-maker can sample the system, can use the system over a period of  $n$  transitions, or can make a terminal decision. If the system is sampled the consecutive sampling rule is operative and if the system is used the  $n$ -step sampling rule is operative. Thus, we shall assume that the prior distribution function of  $\tilde{\underline{P}}$  is  $H(\underline{P}|\gamma) \in \mathcal{H}$ , a family of distributions closed under the  $n$ -step sampling rule. Theorem 2.3.4 implies that  $\mathcal{H}$  is the mixed extension of a family of distributions which is closed under the consecutive sampling rule and, therefore,  $\mathcal{H}$  is also closed under consecutive sampling by Theorem 2.3.3. Thus, if the decision-maker is in the state  $(i, \gamma)$  and chooses either to sample or to use the process, the posterior distribution of  $\tilde{\underline{P}}$  will be a member of  $\mathcal{H}$ .

If it is decided to sample when in state  $(i, \gamma)$ , the supremum of the prior expected reward is given by equation (5.1.4). Suppose, on the other hand, it is decided to use the process under policy  $\underline{\Sigma}$  for  $n > 1$  transitions. The probability that the system will be observed in state  $j$ , unconditional with regard to the prior distribution of  $\tilde{\underline{P}}$ , given that the system starts



in the generalised state  $(1, \gamma)$  and that  $n$  transitions are to be observed under the policy  $\underline{\Sigma}$ , is

$$\bar{p}_{ij}^{(n)}(\underline{\Sigma}, \gamma) = \int_{\underline{\Omega}_n} p_{ij}^{(n)} dF_{\underline{\Sigma}}(\underline{\omega} | \gamma), \quad \begin{matrix} i, j = 1, \dots, N \\ n = 2, 3, \dots \\ \underline{\Sigma} \in \underline{\Sigma} \\ \gamma \in \Gamma \end{matrix} \quad (5.3.1)$$

the  $(i, j)$ th element of the prior expected  $n$ -step transition probability matrix under policy  $\underline{\Sigma}$ . Let  $\bar{q}_i^{(n)}(\underline{\Sigma}, \beta, \gamma)$  denote the prior expected discounted reward earned over  $n$  transitions under the policy  $\underline{\Sigma}$  when the system starts from  $(i, \gamma)$ . Both  $\bar{p}_{ij}^{(n)}(\underline{\Sigma}, \gamma)$  and  $\bar{q}_i^{(n)}(\underline{\Sigma}, \beta, \gamma)$  are discussed in Section 4.1. Let  $T_{ij}(n, \underline{\Sigma}, \gamma)$  denote the parameter of the posterior distribution of  $\tilde{\underline{\Omega}}$  when the system starts from  $(i, \gamma)$  and is observed in state  $j$  after  $n$  transitions under the policy  $\underline{\Sigma}$ . The prior expected reward under these conditions is

$$\bar{q}_i^{(n)}(\underline{\Sigma}, \beta, \gamma) + \beta^n \sum_{j=1}^N \bar{p}_{ij}^{(n)}(\underline{\Sigma}, \gamma) [v_j(T_{ij}(n, \underline{\Sigma}, \gamma)) - e_j] \quad \begin{matrix} i = 1, \dots, N \\ n = 2, 3, \dots \\ \underline{\Sigma} \in \underline{\Sigma} \\ \gamma \in \Gamma \end{matrix} \quad (5.3.2)$$

and, if it is decided to use the system, the supremum of the expected discounted reward is

$$\max_{\underline{\Sigma} \in \underline{\Sigma}} \sup_{n=2, 3, \dots} \left\{ \bar{q}_i^{(n)}(\underline{\Sigma}, \beta, \gamma) + \beta^n \sum_{j=1}^N \bar{p}_{ij}^{(n)}(\underline{\Sigma}, \gamma) [v_j(T_{ij}(n, \underline{\Sigma}, \gamma)) - e_j] \right\} \quad \begin{matrix} i = 1, \dots, N \\ \gamma \in \Gamma \end{matrix} \quad (5.3.3)$$

Finally, if it is decided to make a terminal decision when in the generalized state  $(1, \gamma)$ , the supremum of the expected total discounted reward is

$$\max_{\underline{\Sigma} \in \underline{\Sigma}} \left\{ \bar{v}_i(\underline{\Sigma}, \gamma) \right\}. \quad \begin{matrix} i = 1, \dots, N \\ \gamma \in \Gamma \end{matrix} \quad (5.3.4)$$



We shall anticipate Theorem 5.3.1, which establishes the existence of an optimal sampling strategy for Model II, and write  $\max_{n=2,3,\dots}$  for  $\sup_{n=2,3,\dots}$  in equation (5.3.3). Then the vector function  $\underline{v}(\psi) = (v_1(\psi), \dots, v_N(\psi))$  must satisfy the following functional equation:

$$\boxed{\begin{aligned} v_1(\psi) &= \max \left\{ \max_{1 \leq k \leq K_1} \left\{ \bar{c}_1^k(\psi) - \beta \bar{e}_1^k(\psi) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\psi) v_j(T_{1j}^k(\psi)) \right\} \right. \\ &\quad \left. \max_{\underline{\sigma} \in \Sigma} \max_{n=2,3,\dots} \left\{ \bar{q}_1^{(n)}(\underline{\sigma}, \beta, \psi) + \beta^n \sum_{j=1}^N \bar{p}_{1j}^{(n)}(\underline{\sigma}, \psi) \right. \right. \\ &\quad \left. \left. \times [v_j(T_{1j}^{(n)}(\underline{\sigma}, \psi)) - c_j] \right\} \right. \\ &\quad \left. \max_{\underline{\sigma} \in \Sigma} \left\{ \bar{v}_1(\underline{\sigma}, \psi) \right\} \right\} \end{aligned}}$$

$$i=1, \dots, N \quad (5.3.5)$$

$\psi \in \Psi$

We shall now consider some properties of Model II. These properties, for the most part, parallel those of Model I and the proofs are quite similar to those of Section 5.2. In the following two theorems it is shown that an optimal sampling strategy exists for Model II and that, in an optimal sampling strategy, a terminal decision point is reached with probability one. We then demonstrate the existence of a unique bounded solution to equation (5.3.5) and consider a method of successive approximations, together with a bound on the error of the  $n$ th approximation.

**Theorem 5.3.1** Let  $v_i(\psi, d)$  be the expected total discounted reward in Model II when the system starts in the generalized state  $(i, \psi)$  and the sampling strategy  $d \in D_i$  is used. Let

$$v_i(\psi) = \sup_{d \in D_i} \left\{ v_i(\psi, d) \right\} . \quad i=1, \dots, N \quad (5.3.6)$$

$\psi \in \Psi$



Then there is a sampling strategy  $d^* \in D_i$  such that

$$v_i(\tau) = v_i(\tau, d^*). \quad i=1, \dots, N \quad (\text{5.3.7})$$

$$\tau \in \Psi$$

Proof. Let  $\tilde{D}_i$  be the set of all possible sampling strategies  $\tilde{d}$ , in the adaptive control problem. Then  $D_i \subset \tilde{D}_i$ . Suppose  $\tilde{d} \in \tilde{D}_i$ . Then  $\tilde{d}$  either prescribes a fixed policy,  $\underline{\Sigma}$ , for use on every transition  $n > n^*$ , for some integer  $n^*$ , or else not. In either case,  $\tilde{d}$  is a possible strategy for Model II and  $\tilde{D}_i \subset D_i$ . Thus,  $D_i = \tilde{D}_i$ . The remainder of the proof is analogous to the proof of Theorem 3.1.1. Q.E.D.

**Theorem 5.3.2** In Model II, if the true state of nature  $\underline{Q}$  is a positive matrix and if the terminal functions  $\tilde{V}_i(\underline{\Sigma}, \tau)$  are continuous in  $\tau$  ( $i=1, \dots, N$ ;  $\underline{\Sigma} \in \Sigma$ ), then, with probability one, a terminal decision point is reached in an optimal sampling strategy.

Proof. Assume there is an optimal sampling strategy in which a terminal decision is never made. We shall show a contradiction. Let a decision point in the sample history be a point in time at which the state of the system is made known to the decision-maker. Then the assumption is that there are an infinite number of decision points. There is at least one state,  $i$ , which is observed infinitely often and at least one policy,  $\underline{\Sigma}$ , and transition interval,  $n$ , which are used infinitely often in state  $i$ . Lemma 2.3.7 and the positivity of  $\underline{Q}$  imply that every state is observed infinitely often with probability one. Since a terminal decision is never made, there is a finite integer,  $\mu$ , such that, if  $n$  is a transition interval which is used in the sampling strategy, then  $n \leq \mu$ . For if not, an infinite transition interval is used at some stage, which is equivalent to a terminal decision. Thus, there is a finite set of ordered pairs,



$(n, \underline{\sigma})$ , where  $n \in \{1, 2, \dots, \mu\}$  and  $\underline{\sigma} \in \Sigma$ , which describe the decisions made at each decision point. We may assume, without loss of generality, that all members of this finite set are used infinitely often in the sampling strategy, since any pair,  $(n, \underline{\sigma})$ , which is used only a finite number of times is eventually dominated. The conditions of Theorem 2.3.9 are satisfied, and, therefore, the mass of the posterior distribution of tends, with probability one, to concentrate at  $\underline{\Omega}$  as  $v$ , the number of decision points, goes to infinity. If  $H(\underline{\sigma} | \Psi^*)$  is the distribution function which places the unit mass of probability on  $\underline{\Omega}$ , then  $\Psi \rightarrow \Psi^*$  as  $v \rightarrow \infty$  and equation (5.3.5) becomes<sup>#</sup>

$$v_1(\Psi^*) = \max_{\underline{\sigma} \in \Sigma} \left[ \begin{aligned} & \max_{1 \leq k \leq K_1} \left\{ \bar{q}_1^k(\Psi^*) - \beta \bar{c}_1^k(\Psi^*) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\Psi^*) v_j(\Psi^*) \right\} \\ & \max_{n \in \Sigma} \max_{n=2, \dots, \mu} \left\{ \bar{q}_1^{(n)}(\underline{\sigma}, \beta, \Psi^*) + \beta^n \sum_{j=1}^N \bar{p}_{1j}^{(n)}(\underline{\sigma}, \Psi^*) \right. \\ & \quad \times [v_j(\Psi^*) - c_j] \Big\} \\ & \max_{\underline{\sigma} \in \Sigma} \left\{ v_i(\underline{\sigma}, \Psi^*) \right\} \end{aligned} \right]_{i=1, \dots, N} \quad (5.3.8)$$

It was shown, during the proof of Theorem 5.1.2, that if

$$v_1(\Psi^*) = \max_{1 \leq k \leq K_1} \left\{ \bar{q}_1^k(\Psi^*) - \beta \bar{c}_1^k(\Psi^*) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\Psi^*) v_j(\Psi^*) \right\},$$

$$i=1, \dots, N \quad (5.3.9)$$

a contradiction results. Suppose, then, that

$$v_1(\Psi^*) = \max_{\underline{\sigma} \in \Sigma} \max_{n=2, \dots, \mu} \left\{ \bar{q}_1^{(n)}(\underline{\sigma}, \beta, \Psi^*) + \beta^n \sum_{j=1}^N \bar{p}_{1j}^{(n)}(\underline{\sigma}, \Psi^*) [v_j(\Psi^*) - c_j] \right\},$$

$$i=1, \dots, N \quad (5.3.10)$$

<sup>#</sup> The argument remains valid when the distribution which places the unit mass of probability on  $\underline{\Omega}$  is not a member of  $\Psi$ .



We may construct a new set of policies as follows. Let  $\underline{\gamma} = (\underline{s}_1, \dots, \underline{s}_N)$  be a policy vector, where  $\underline{s}_k = (n_k, \Sigma_k)$  is a choice of a transition interval,  $n_k \in \{1, \dots, \mu\}$ , and a policy,  $\Sigma_k \in \Sigma$ . If the alternative  $s_k$  is selected in state  $i$ , then the system goes to state  $j$  with probability  $\tilde{p}_{ij}^{(n_k)}(\Sigma_k, \psi^*)$ , earning the expected reward  $\tilde{q}_j^{(n_k)}(\Sigma_k, \beta, \psi^*) = \beta^{n_k} c_j$ . If  $S$  is the set of all possible alternatives  $s_k$ , then equation (5.3.10) can be written as

$$v_i(\psi^*) = \max_{s_k \in S} \left\{ \sum_{j=1}^N \tilde{p}_{ij}^{(n_k)}(\Sigma_k, \psi^*) [\tilde{q}_j^{(n_k)}(\Sigma_k, \beta, \psi^*) - \beta^{n_k} c_j + \beta^{n_k} v_j(\psi^*)] \right\}. \\ i=1, \dots, N \quad (5.3.11)$$

Equation (5.3.11) has the same formal structure as equation (5.1.11) in the proof of Theorem 5.1.2. An argument similar to that leading to equation (5.1.13) shows the contradiction

$$v_i(\psi^*) < \max_{\Sigma \in \Sigma} \left\{ \tilde{v}_i(\Sigma, \psi^*) \right\}. \quad i=1, \dots, N \quad (5.3.12)$$

Thus, with probability one, a terminal decision point is reached. Q.E.D.

Let us now consider the existence and uniqueness of solutions to the functional equation (5.3.5). Let the sequence of vector functions,

$\underline{x}(n, \psi)$ , where  $\underline{x}(n, \psi) = (v_1(n, \psi), \dots, v_N(n, \psi))$ , be defined by the following equations:

$$\begin{aligned} v_1(n+1, \psi) &= \max_{1 \leq k \leq \bar{k}} \left\{ \tilde{q}_1^k(\psi) - \beta \tilde{c}_1^k(\psi) + \beta \sum_{j=1}^N \tilde{p}_{1j}^k(\psi) v_j(n, \tau_{1j}^k(\psi)) \right\} \\ &\quad \boxed{\max_{\Sigma \in \Sigma} \max_{n=2, 3, \dots, n+1} \left\{ \tilde{q}_1^{(v)}(\Sigma, \beta, \psi) + \beta^v \sum_{j=1}^N \tilde{p}_{1j}^{(v)}(\Sigma, \psi) \right. \\ &\quad \quad \quad \left. \times [v_j(n, \tau_{1j}^{(v, \Sigma, \psi)}) - c_j] \right\} \\ &\quad \max_{\Sigma \in \Sigma} \left\{ \tilde{v}_1(\Sigma, \psi) \right\} \end{aligned}$$

$$\begin{aligned} &i=1, \dots, N \quad (5.3.13a) \\ &n=0, 1, 2, \dots \\ &\psi \in \Psi \\ &0 \leq \beta < 1 \end{aligned}$$



$$v_1(0, \psi) = \max_{\substack{i \in \Sigma \\ \psi \in \Psi}} \left\{ \bar{v}_1(\underline{\sigma}, \psi) \right\} \quad . \quad i=1, \dots, N \quad (5.3.13b)$$

**Lemma 5.3.3** If  $R^* = \max_{i,j,k} \left\{ \left| \frac{k}{r_{ij}} \right| \right\}$  and  $C = \max_j \left\{ c_j \right\}$ , then the functions  $v_1(n, \psi)$  defined by (5.3.13) have the bound

$$|v_1(n, \psi)| \leq \frac{R^* + \beta C}{1-\beta} . \quad \begin{array}{l} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \psi \in \Psi \\ 0 \leq \beta < 1 \end{array} \quad (5.3.14)$$

**Proof.** The proof is by induction. Equation (5.2.4) shows that (5.3.14) holds for  $n=0$ . Assume it holds for  $n$ . Since  $\left| \frac{-1}{q_1}(\underline{\sigma}, \beta, \psi) \right| = \left| \frac{\sigma^{-1}}{q_1}(\psi) \right| \leq R^*$ , an induction using equation (4.1.9) shows that

$$\left| \frac{-1}{q_1}(\underline{\sigma}, \beta, \psi) \right| \leq \frac{1-\beta^\nu}{1-\beta} R^* . \quad \begin{array}{l} i=1, \dots, N \\ \nu=1, 2, 3, \dots \\ \frac{\sigma}{\Gamma} \in \Sigma \\ \Gamma \in \Psi \end{array} \quad (5.3.15)$$

Thus, from (5.3.13a) and the induction hypothesis, we have

$$\begin{aligned} |v_1(n+1, \psi)| &\leq \max \left[ R^* + \beta C + \beta \frac{R^* + \beta C}{1-\beta}, \right. \\ &\quad \left. \max_{\nu=2, 3, \dots, n+1} \left\{ \frac{1-\beta^\nu}{1-\beta} R^* + \beta^\nu C + \beta^\nu \frac{R^* + \beta C}{1-\beta} \right\}, \right. \\ &\quad \left. \frac{R^*}{1-\beta} \right] \\ &= \max \left[ \frac{R^* + \beta C}{1-\beta}, \max_{\nu=2, \dots, n+1} \left\{ \frac{R^* + \beta^\nu C}{1-\beta} \right\} \right] \\ &= \frac{R^* + \beta C}{1-\beta}. \end{aligned} \quad (5.3.16)$$

Q.E.D.

**Theorem 5.3.4** If  $y(n, \psi)$  is defined by equation (5.3.13), then  $\{v_1(n, \psi)\}$  is a monotone increasing sequence ( $i=1, \dots, N$ ) and  $\lim_{n \rightarrow \infty} v_1(n, \psi)$  exists and is a solution of (5.3.5).



**PROOF.** The proof that  $\{v_i(n, \tau)\}$  is monotone increasing is inductive. Clearly,  $v_i(1, \tau) \geq v_i(0, \tau)$  for  $i=1, \dots, N$  and  $\tau \in \mathbb{T}$ . Assume that  $v_i(n, \tau) \geq v_i(n-1, \tau)$ . If, for some  $\underline{\sigma} \in \Sigma$  and some integer  $v$ ,

$$v_i(n, \tau) = q_i^{-(v)}(\underline{\sigma}, \beta, \tau) + \beta^v \sum_{j=1}^N p_{ij}^{-(v)}(\underline{\sigma}, \tau) [v_j(n-1, T_{1,j}(v_i \underline{\sigma} + \tau)) - e_j] \quad (5.3.17)$$

then

$$v_i(n+1, \tau) - v_i(n, \tau) \geq \beta^v \sum_{j=1}^N p_{ij}^{-(v)}(\underline{\sigma}, \tau) [v_j(n, T_{1,j}(v_i \underline{\sigma} + \tau)) - v_j(n-1, T_{1,j}(v_i \underline{\sigma} + \tau))] \\ \geq 0. \quad (5.3.18)$$

If  $v_i(n, \tau) = \max_{\underline{\sigma} \in \Sigma} \{\bar{v}_i(\underline{\sigma}, \tau)\}$ , then  $v_i(n+1, \tau) \geq v_i(n, \tau)$ , and, using (5.2.8), we have, in all cases,

$$v_i(n+1, \tau) \geq v_i(n, \tau), \quad \begin{matrix} i=1, \dots, N \\ \tau \in \mathbb{T} \end{matrix} \quad (5.3.19)$$

proving monotonicity. By Lemma 5.3.3, the sequence  $\{v_i(n, \tau)\}$  is bounded and, therefore,  $\lim_{n \rightarrow \infty} v_i(n, \tau)$  exists. That the limit satisfies (5.3.5) follows by letting  $n \rightarrow \infty$  in (5.3.13a). Q.E.D.

The remaining theorems, the proofs of which parallel very closely those of corresponding theorems in Section 5.2, are stated without proof.

**Theorem 5.3.5** There is a unique bounded solution to equation (5.3.5).

**Theorem 5.3.6** Let the error of the  $n$ th approximant,  $v_i(n, \tau)$ , be defined as

$$e_i(n, \tau) = v_i(\tau) - v_i(n, \tau), \quad (5.3.20)$$

$i=1, \dots, N$   
 $n=0, 1, 2, \dots$   
 $\tau \in \mathbb{T}$

where  $\bar{v}(\tau) = (v_1(\tau), \dots, v_N(\tau))$  is the unique bounded solution of



equation (5.3.5) and  $y(n, \gamma)$  is defined by (5.3.13). Then  $e_1(n, \gamma)$  has the bounds

$$0 \leq e_1(n, \gamma) \leq \beta^n \frac{R \cdot r}{1-\beta}, \quad \begin{aligned} i=1, \dots, N \\ n=0, 1, 2, \dots \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{aligned} \quad (5.3.21)$$

where  $R$  and  $r$  are defined by (3.3.3).

Corollary 5.3.2 If  $y(n, \gamma)$  is defined by equation (5.3.13) and  $y(\gamma)$  is the unique bounded solution of (5.3.5), then  $\{y(n, \gamma)\} \rightarrow y(\gamma)$  uniformly in  $\gamma$ .

Theorem 5.3.3 If the prior distribution function of  $\underline{\Phi}$  is  $N(\underline{\Phi} | \gamma) \subset \mathcal{N}$ , a family of distributions continuous in  $\gamma$  which is closed under v-step sampling, then  $y(\gamma)$ , the unique bounded solution of (5.3.5), is continuous in  $\gamma$ .

The numerical solution of Model II involves considerably more computation than does the solution of Model I. Not only does equation (5.3.13), the successive approximation scheme for Model II, involve evaluation of more terms than does the corresponding scheme for Model I, but the requirement that  $\mathcal{N}$ , the family of prior distributions of  $\underline{\Phi}$ , be closed under v-step sampling implies that  $\mathcal{N}$  is the mixed extension of a family of distributions closed under consecutive sampling. This means that the number of parameters which must be handled in computing solutions to Model II is larger than that required for solving Model I. The additional complexity of Model II is probably worthwhile only in the case of a prior distribution which is tight on some rows of  $\underline{\Phi}$  and loose on others and where the cost of sampling is high.

We note that, while the aim of Model II is to allow the decision-maker



to sample only those states in which there are transition probability vectors with loose marginal prior distributions, he does not have full control over the future states in which the system may be observed. For example, suppose it is desired to sample the system only when it is in state  $i$ . Then a sampling strategy must be chosen which trades off the expected discounted earnings of the system against the need for a high probability that the system enters state  $i$  at each decision point.

We remark, in this connection, that a decision to use the process when in state  $i$  does not necessarily imply that the consecutive sampling alternative is dominated at future decision points when the system is found in state  $i$ . Such dominance may hold under a sample history which reduces the marginal variances of the alternative transition probabilities in the  $i$ th state, but there is certainly no reason to expect this to be the case under a sequence of observations which increases the marginal variances of some of these transition probabilities.

#### 5.4 Approximate Terminal Decisions.

In Models I and II it is necessary to evaluate expressions of the form

$$\tilde{V}_i(\Sigma^*, \tau) = \max_{\Sigma \in \Sigma} \left\{ \tilde{V}_i(\Sigma, \tau) \right\}, \quad (5.4.1)$$

where  $\tilde{V}_i(\Sigma, \tau)$  is the expected total discounted reward earned over an infinite period under the policy  $\Sigma$  when the system starts from the generalized state  $(i, \tau)$ . Since  $\Sigma$  may contain a large number of policies, it is desirable to find methods of solving (5.4.1) for the maximizing policy  $\Sigma^*$  which avoid a direct search over all elements of  $\Sigma$ . This problem has not been solved, but some preliminary remarks concerning the approximation of  $\Sigma^*$  are offered in this section. It will be seen that these remarks



are also applicable to the problem of selecting a policy which maximizes the expected gain,  $\tilde{g}(\underline{\Sigma}, \gamma)$ , which was discussed in Section 4.4.

Let  $v_1(\underline{\Sigma}, \underline{\Phi})$  be the conditional expected total discounted reward over an infinite period under policy  $\underline{\Sigma}$  when the system starts from state 1 and  $\tilde{\underline{\Phi}} = \underline{\Phi}$ . The policy which maximizes this reward can be found efficiently by means of Howard's policy iteration algorithm [22]. It was seen in the proofs of Theorems 5.1.2 and 5.3.2 that, as the number of observations in Model I or Model II goes to infinity, the mass of the posterior probability distribution of  $\tilde{\underline{\Phi}}$  tends, with probability one, to concentrate at the true state of nature. Thus, if  $\bar{\underline{\Phi}}$  is the mean of the distribution of  $\tilde{\underline{\Phi}}$ , we can approximate  $\underline{\Sigma}^*$  by  $\hat{\underline{\Sigma}}$ , where  $\hat{\underline{\Sigma}}$  is defined by

$$v_1(\hat{\underline{\Sigma}}, \bar{\underline{\Phi}}) = \max_{\underline{\Sigma} \in \Sigma} \{v_1(\underline{\Sigma}, \bar{\underline{\Phi}})\}. \quad (5.4.2)$$

The error of this approximation goes to zero with probability one as the number of observations of the process goes to infinity. We consider here a bound on the error,

$$\phi_1(\hat{\underline{\Sigma}}, \gamma) = \tilde{v}_1(\underline{\Sigma}^*, \gamma) - \tilde{v}_1(\hat{\underline{\Sigma}}, \gamma). \quad (5.4.3)$$

Let the policies,  $\underline{\Sigma}_j$ , be indexed by  $j$ . Thus,  $\Sigma = \{\underline{\Sigma}_1, \dots, \underline{\Sigma}_J\}$ , where  $J$  is the number of distinct policies in  $\Sigma$ . For a fixed index  $i$ , let  $\mathcal{S}_{K,N}$  be partitioned into a set of  $J$  mutually exclusive and exhaustive subsets,  $S_{ij}$ , such that, if  $\underline{\Phi} \in S_{ij}$ , then

$$v_1(\underline{\Sigma}_j, \underline{\Phi}) = \max_{\underline{\Sigma} \in \Sigma} \{v_1(\underline{\Sigma}, \underline{\Phi})\}. \quad (5.4.4)$$

If  $h(\underline{\Phi} | \gamma)$  is the prior distribution function of  $\tilde{\underline{\Phi}}$ , let

$$P_{ij}(\gamma) = \int_{S_{ij}} h(\underline{\Phi} | \gamma) \quad j=1, \dots, J \quad \gamma \in \mathbb{R} \quad (5.4.5)$$

denote the prior probability that  $\tilde{\underline{\Phi}}$  belongs to  $S_{ij}$ . Since the sets  $S_{ij}$  partition  $\mathcal{S}_{K,N}$ ,



-140-

$$\sum_{j=1}^J p_{ij}(\gamma) = 1. \quad \begin{matrix} \gamma \in \Gamma \\ i=1, \dots, N \end{matrix} \quad (5.4.6)$$

Let  $\tilde{v}_{ij}(\underline{\sigma}, \gamma)$  be the conditional expected value of  $v_i(\underline{\sigma}, \tilde{\gamma})$ , given that  $\underline{\sigma} \in S_{ij}$ :

$$\tilde{v}_{ij}(\underline{\sigma}, \gamma) = \frac{1}{P_{ij}(\gamma)} \int_{S_{ij}} v_i(\underline{\sigma}, \underline{\gamma}) dH(\underline{\gamma} | \gamma). \quad \begin{matrix} j=1, \dots, J \\ \underline{\sigma} \in \Sigma \\ \gamma \in \Gamma \end{matrix} \quad (5.4.7)$$

We note that (5.4.4) implies

$$\tilde{v}_{ij}(\underline{\sigma}_j, \gamma) \geq \tilde{v}_{ij}(\underline{\sigma}, \gamma). \quad \begin{matrix} j=1, \dots, J \\ \underline{\sigma} \in \Sigma \\ \gamma \in \Gamma \end{matrix} \quad (5.4.8)$$

Let  $R(\underline{\sigma})$  and  $r(\underline{\sigma})$  be the maximum and minimum transition rewards when the policy  $\underline{\sigma} = (\sigma_1, \dots, \sigma_N)$  is used,

$$R(\underline{\sigma}) = \max_{i,j} \left\{ r_{ij}^{\sigma_i} \right\} \quad \underline{\sigma} \in \Sigma \quad (5.4.9a)$$

$$r(\underline{\sigma}) = \min_{i,j} \left\{ r_{ij}^{\sigma_i} \right\}, \quad \underline{\sigma} \in \Sigma \quad (5.4.9b)$$

and let  $r$  and  $R$  be defined by equation (3.3.3). By equation (4.3.1),

$$v_i(\underline{\sigma}, \underline{\gamma}) = \sum_{n=0}^{\infty} \beta^n \sum_{j=1}^N \sum_{k=1}^N p_{ij}^{(n)}(\underline{\sigma}) p_{jk}^{\sigma_j} r_{jk}^{\sigma_k} \quad \begin{matrix} \underline{\sigma} \in \Sigma \\ \underline{\gamma} \in \Gamma_{K,N} \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.10)$$

$$\begin{matrix} \underline{\sigma} \in \Sigma \\ \underline{\gamma} \in \Gamma_{K,N} \\ 0 \leq \beta < 1 \end{matrix}$$

and, therefore,

$$\frac{R}{1-\beta} \leq \frac{v_i(\underline{\sigma})}{1-\beta} \leq \tilde{v}_{ij}(\underline{\sigma}, \gamma) \leq \frac{R(\underline{\sigma})}{1-\beta} \leq \frac{R}{1-\beta} \quad \begin{matrix} \underline{\sigma} \in \Sigma \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.11)$$



$$\frac{r(\sigma)}{1-\beta} \leq \bar{v}_{1j}(\underline{\sigma}, \tau) \leq \frac{R(\sigma)}{1-\beta} \quad \begin{matrix} j=1, \dots, J \\ \underline{\sigma} \in \Sigma \\ \tau \in \Psi \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.12)$$

Lemma 5.4.1 For  $j=1, \dots, J$ , the following inequality is valid,

$$\bar{v}_{1j}(\underline{\sigma}_j, \tau) P_{1j}(\tau) \leq \bar{v}_1(\underline{\sigma}_j, \tau) - (1 - P_{1j}(\tau)) \frac{r(\sigma_j)}{1-\beta} \quad \begin{matrix} j=1, \dots, J \\ \tau \in \Psi \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.13)$$

Proof. Since the sets  $S_{1j}$  partition  $\mathcal{S}_{K, N}$  we have, using (5.4.12)

$$\begin{aligned} \bar{v}_1(\underline{\sigma}_j, \tau) &= \bar{v}_{1j}(\underline{\sigma}_j, \tau) P_{1j}(\tau) + \sum_{k \neq j} \bar{v}_{1k}(\underline{\sigma}_j, \tau) P_{1k}(\tau) \\ &\geq \bar{v}_{1j}(\underline{\sigma}_j, \tau) P_{1j}(\tau) + (1 - P_{1j}(\tau)) \frac{r(\sigma_j)}{1-\beta}. \end{aligned} \quad (5.4.14)$$

Equation (5.4.13) is a rearrangement of (5.4.14). Q.E.D.

Lemma 5.4.2 For any policy  $\underline{\sigma} \in \Sigma$  and any index  $j \in \{1, \dots, J\}$ , the expected discounted reward under the policy  $\underline{\sigma}$  has the upper bound

$$\bar{v}_1(\underline{\sigma}, \tau) \leq \bar{v}_2(\underline{\sigma}_j, \tau) + \frac{1 - P_{1j}(\tau)}{1-\beta} (R - r(\sigma_j)). \quad \begin{matrix} j=1, \dots, J \\ \underline{\sigma} \in \Sigma \\ \tau \in \Psi \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.15)$$

Proof. We have, using equations (5.4.11), (5.4.8), and Lemma 5.4.1,

$$\begin{aligned} \bar{v}_1(\underline{\sigma}, \tau) &= \bar{v}_{1j}(\underline{\sigma}, \tau) P_{1j}(\tau) + \sum_{k \neq j} \bar{v}_{1k}(\underline{\sigma}, \tau) P_{1k}(\tau) \\ &\leq \bar{v}_{1j}(\underline{\sigma}, \tau) P_{1j}(\tau) + (1 - P_{1j}(\tau)) \frac{R}{1-\beta} \end{aligned}$$



$$\begin{aligned} &\leq \tilde{V}_{1j}(\underline{\Sigma}_j, \gamma) P_{1j}(\gamma) + (1-P_{1j}(\gamma)) \frac{R}{1-\beta} \\ &\leq \tilde{V}_2(\underline{\Sigma}_j, \gamma) + \frac{1-P_{1j}(\gamma)}{1-\beta} (R+\rho(\underline{\Sigma}_j)). \end{aligned} \quad (5.4.16)$$

Q.E.D.

Theorem 5.4.3 Let  $e_1(\underline{\Sigma}_j, \gamma)$  be defined by

$$e_1(\underline{\Sigma}_j, \gamma) = \tilde{V}_1(\underline{\Sigma}^*, \gamma) - \tilde{V}_1(\underline{\Sigma}_j, \gamma) \quad (5.4.17)$$

$\begin{matrix} j=1, \dots, J \\ i=1, \dots, N \\ \gamma \in \Gamma \end{matrix}$

where  $\underline{\Sigma}^*$  is the maximizing terminal policy defined by equation (5.4.1).

Then  $e_1(\underline{\Sigma}_j, \gamma)$  has the bounds

$$0 \leq e_1(\underline{\Sigma}_j, \gamma) \leq \frac{1-P_{1j}(\gamma)}{1-\beta} (R+\rho(\underline{\Sigma}_j)). \quad \begin{matrix} i=1, \dots, N \\ j=1, \dots, J \\ \gamma \in \Gamma \\ 0 \leq \beta < 1 \end{matrix} \quad (5.4.18)$$

Proof. By equation (5.4.1),  $\tilde{V}_1(\underline{\Sigma}^*, \gamma) \geq \tilde{V}_1(\underline{\Sigma}_j, \gamma)$  and  $e_1(\underline{\Sigma}_j, \gamma) \geq 0$ .

The upper half of the inequality follows from (5.4.15). Q.E.D.

In order to bound  $e_1(\underline{\Sigma}_j, \gamma)$  using equation (5.4.18), it is necessary to evaluate  $P_{1j}(\gamma)$ . This problem has not been completely solved, chiefly because there is no satisfactory method of finding the boundaries of the set  $S_{1j}$ . Moreover,  $S_{1j}$  is not necessarily a connected set, which further complicates the problem. The probability  $P_{1j}(\gamma)$  can be estimated by using numerical or Monte Carlo techniques.

If  $g(\underline{\Sigma}, \widetilde{\underline{\Phi}})$  is the gain of a Markov chain with alternatives when operated indefinitely under the policy  $\underline{\Sigma}$ , and if  $\widetilde{g}(\underline{\Sigma}, \gamma)$  is the corresponding expected value when  $\widetilde{\underline{\Phi}}$  has the distribution function  $\pi(\underline{\Omega}|\gamma)$ , then, as we shall see in the next section, it is often necessary to evaluate expressions of the form



$$\tilde{g}(\underline{\Sigma}^*, \psi) = \max_{\underline{\Sigma} \in \Sigma} \{ \tilde{g}(\underline{\Sigma}, \psi) \}. \quad (5.4.19)$$

If  $\bar{P}$  is the mean of the distribution  $H(P|\psi)$  we may wish to approximate  $\underline{\Sigma}^*$  by  $\hat{\underline{\Sigma}}$ , defined by the expression

$$g(\hat{\underline{\Sigma}}, \bar{P}) = \max_{\underline{\Sigma} \in \Sigma} \{ g(\underline{\Sigma}, \bar{P}) \}. \quad (5.4.20)$$

There are efficient algorithms for the solution of (5.4.20) [16, 22]. Let the error of the approximation  $\underline{\Sigma}_j$  be defined as

$$e(\underline{\Sigma}_j, \psi) = \tilde{g}(\underline{\Sigma}^*, \psi) - \tilde{g}(\underline{\Sigma}_j, \psi). \quad j=1, \dots, J \quad \forall \psi \in \Psi \quad (5.4.21)$$

A bound on  $e(\underline{\Sigma}_j, \psi)$  similar to that of equation (5.4.18) is easily derived.

Let  $\mathcal{S}_{K,N}^*$  be the set of all positive  $K \times N$  generalized stochastic matrices and let  $\mathcal{S}_{K,N}^*$  be partitioned into  $J$  sets,  $S_j$ , where, if  $\bar{P} \in S_j$ , then

$$g(\underline{\Sigma}_j, \bar{P}) = \max_{\underline{\Sigma} \in \Sigma} \{ g(\underline{\Sigma}, \bar{P}) \}. \quad (5.4.22)$$

If  $H(P|\psi)$  is the prior distribution function of  $\bar{P}$ , let

$$p_j(\psi) = \int_{S_j} dH(\bar{P}|\psi) \quad j=1, \dots, J \quad \forall \psi \in \Psi \quad (5.4.23)$$

be the prior probability that  $\bar{P} \in S_j$ . Let

$$\tilde{g}_j(\underline{\Sigma}, \psi) = \frac{1}{p_j(\psi)} \int_{S_j} g(\underline{\Sigma}, \bar{P}) dH(\bar{P}|\psi) \quad j=1, \dots, J \quad \begin{matrix} \underline{\Sigma} \in \Sigma \\ \bar{P} \in \Psi \end{matrix} \quad (5.4.24)$$

be the conditional expectation of  $g(\underline{\Sigma}, \bar{P})$  given that  $\bar{P} \in S_j$ . Then, by (5.4.22),

$$\tilde{g}_j(\underline{\Sigma}_j, \psi) \geq \tilde{g}_j(\underline{\Sigma}, \psi). \quad j=1, \dots, J \quad \begin{matrix} \underline{\Sigma} \in \Sigma \\ \bar{P} \in \Psi \end{matrix} \quad (5.4.25)$$

If  $R(\underline{\Sigma})$  and  $r(\underline{\Sigma})$  are defined by (5.4.9) and  $R$  and  $r$  are defined by (3.3.3), (4.4.1) implies the inequalities

$$r \leq \tilde{g}(\underline{\Sigma}, \psi) \leq R \quad \underline{\Sigma} \in \Sigma \quad (5.4.26)$$

$$r(\underline{\Sigma}) \leq \tilde{g}_j(\underline{\Sigma}, \psi) \leq R(\underline{\Sigma}). \quad j=1, \dots, J \quad \begin{matrix} \underline{\Sigma} \in \Sigma \\ \bar{P} \in \Psi \end{matrix} \quad (5.4.27)$$



Lemma 5.4.4 For  $j=1, \dots, J$

$$\tilde{g}_j(\underline{\pi}_j, \gamma) P_j(\gamma) \leq \tilde{g}(\underline{\pi}_j, \gamma) = (1-P_j(\gamma)) r(\underline{\pi}_j). \quad (5.4.28)$$

$\frac{j=1, \dots, J}{\gamma \in \Gamma}$

Proof. Using (5.4.27),

$$\begin{aligned} \tilde{g}(\underline{\pi}, \gamma) &= \tilde{g}_j(\underline{\pi}_j, \gamma) P_j(\gamma) + \sum_{k \neq j} \tilde{g}_k(\underline{\pi}_j, \gamma) P_k(\gamma) \\ &\geq \tilde{g}_j(\underline{\pi}_j, \gamma) P_j(\gamma) + (1-P_j(\gamma)) r(\underline{\pi}_j). \end{aligned} \quad (5.4.29)$$

Q.E.D.

Lemma 5.4.5 For any policy  $\underline{\pi} \in \Sigma$  and any  $j \in \{1, \dots, J\}$ ,

$$\tilde{g}(\underline{\pi}, \gamma) \leq \tilde{g}(\underline{\pi}_j, \gamma) + (1-P_j(\gamma))(R-r(\underline{\pi}_j)). \quad (5.4.30)$$

$\frac{j=1, \dots, J}{\frac{\underline{\pi} \in \Sigma}{\gamma \in \Gamma}}$

Proof. By (5.4.26), (5.4.25), and Lemma 5.4.4,

$$\begin{aligned} \tilde{g}(\underline{\pi}, \gamma) &= \tilde{g}_j(\underline{\pi}, \gamma) P_j(\gamma) + \sum_{k \neq j} \tilde{g}_k(\underline{\pi}, \gamma) P_k(\gamma) \\ &\leq \tilde{g}_j(\underline{\pi}_j, \gamma) P_j(\gamma) + (1-P_j(\gamma)) R \\ &\leq \tilde{g}(\underline{\pi}_j, \gamma) + (1-P_j(\gamma))(R-r(\underline{\pi}_j)). \end{aligned} \quad (5.4.31)$$

Q.E.D.

Theorem 5.4.6 The error function  $e(\underline{\pi}_j, \gamma)$  defined by equation (5.4.21) has the bound

$$0 \leq e(\underline{\pi}_j, \gamma) \leq (1-P_j(\gamma))(R-r(\underline{\pi}_j)). \quad \frac{j=1, \dots, J}{\gamma \in \Gamma} \quad (5.4.32)$$

Proof. The theorem follows directly from equations (5.4.19) and (5.4.30). Q.E.D.

## 5.5 Undiscounted Processes.

We have commented in Section 3.6 on the lack of a clear criterion for



making decisions in an adaptive control model with no discounting. These remarks apply as well to undiscounted terminal control models. One criterion we may use is to consider the class of sampling strategies which maximize the expected steady-state gain of the process, then to choose from this class the strategy which maximizes the expected reward over the transient period which precedes the terminal decision. This criterion is made precise in the present section, where Models III and IV, the undiscounted analogues of Models I and II, are introduced. No analysis of these models has been carried out.

**5.5.1 Model III.** In Model III it is assumed that the process will be sampled consecutively until a terminal decision point is reached, at which time a terminal policy is selected and the system is operated under this policy over a finite terminal operation period.

Let  $v_i(\tau; v)$  be the supremum of the expected reward over a period whose terminal operation phase lasts for  $v$  transitions when the system starts in state  $i$  and the prior distribution of  $\underline{P}$  is  $H(P(\tau)) \in \mathcal{H}$ , a family of distributions closed under consecutive sampling. If, when in state  $(i, \tau)$ , it is decided to sample at least once more, the supremum of the prior expected reward is given by equation (5.1.4) with  $\beta = 1$ .

Since we shall be concerned with large values of  $v$ , we assume that, when it is decided to cease sampling, a terminal policy will be selected which maximizes the steady-state gain of the system  $\bar{g}(\underline{\sigma}, \tau)$ . Therefore, when it is decided to cease sampling, the supremum of the expected reward over the terminal period is

$$\max_{\underline{\sigma} \in \Sigma} \left\{ v_{\bar{g}}(\underline{\sigma}, \tau) \right\}. \quad (5.5.1)$$



Thus, under the assumptions of Model III,  $v_1(\gamma; v)$  must satisfy the following functional equations,

$$v_1(\gamma; v) = \max \left[ \max_{1 \leq k \leq K_1} \left\{ \frac{\alpha^k}{\beta} (\gamma) - \bar{q}_1^k(\gamma) + \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) v_j(\tau_{1j}^k(\gamma); v) \right\} \right]$$

$$\boxed{\max_{\sigma \in \Sigma} \left\{ v_1^\sigma(\sigma; \gamma) \right\}}$$

for  $i = 1, \dots, N$  (5.5.2)  
 $\gamma \in \Gamma$   
 $v = 1, 2, 3, \dots$

The arguments of Sections 5.1 and 5.2 required that the discount factor  $\beta$  be less than unity and, therefore, are not directly applicable to Model III. The existence and properties of solutions to (5.5.2) are matters for future investigation.

Equation (5.5.2) yields the supremum of the expected reward over a period with terminal phase of length  $v$  and also yields a decision for the current transition interval. This decision, which is either the selection of an alternative to be sampled or a terminal policy, will be called a  $v$ -optimal decision. A  $v$ -optimal decision which is the same for all  $v$  sufficiently large will be called an optimal decision. Since, for large  $v$ , every  $v$ -optimal decision maximizes the expected gain and also maximizes the total reward over the sampling period, it is seen that an optimal decision, as defined here, satisfies the criterion set forth at the beginning of the section. The existence and nature of optimal decisions have not yet been investigated.

**5.5.2 Model IV.** We now assume that the decision-maker can sample or use the process at any time prior to the terminal decision, independently of his past decisions. Let  $v_1(\gamma; v)$  be the supremum of the expected reward



over a period with terminal operation phase of length  $v$  when the system starts in state  $i$  with the prior distribution  $H(\underline{\varphi}|\Psi)$ . It is assumed that  $H(\underline{\varphi}|\Psi) \in \mathcal{H}$ , a family of distributions closed under  $v$ -step sampling. Following the arguments of Section 5.3 and of the previous paragraph, it is seen that  $v_i(\Psi; v)$  must satisfy the following functional equation.

$$v_i(\Psi; v) = \max \left[ \begin{array}{l} \max_{1 \leq k \leq K_1} \left\{ \bar{q}_k^*(\Psi) - \bar{q}_k^*(\Psi) + \sum_{j=1}^N \bar{p}_{1,j}^k(\Psi) v_j(T_{1,j}^k(\Psi); v) \right\} \\ \max_{\underline{\sigma} \in \Sigma} \max_{n=2,3,\dots} \left\{ \bar{q}_2^{(n)}(\underline{\sigma}, i, \Psi) + \sum_{j=1}^N \bar{p}_{1,j}^{(n)}(\underline{\sigma}; \Psi) \right. \\ \quad \left. \times [v_j(T_{1,j}^{(n)}(\underline{\sigma}, \Sigma), \Psi) - e_j] \right\} \\ \max_{\underline{\sigma} \in \Sigma} \left\{ \bar{v}_2(\underline{\sigma}, \Psi) \right\} \end{array} \right]$$

$i=1, \dots, N \quad (5.5.3)$   
 $\Psi \in \mathbb{P}$   
 $v=1, 2, 3, \dots$

The remarks made above concerning  $v$ -optimal decisions and optimal decisions apply as well to Model IV.

By approaching optimal decisions for the undiscounted terminal control models by means of  $v$ -optimal decisions we have emphasized the fact that an undiscounted infinite horizon model is an approximation to a system which runs for a long, but finite, period. We can equally well view the undiscounted model as an approximation to a system with a discount factor very close to unity. Thus, another approach to the solution of the undiscounted terminal control problem is to let  $\beta \rightarrow 1$  in Models I and II. The existence and properties of solutions obtained in this manner and their relation to solutions obtained via  $v$ -optimal decisions have not yet been investigated.



### 5.6 Discounted Processes with Set-Up Costs.

In many processes which can be modelled as a Markov chain with alternatives there is a cost associated with changing alternatives in each state. Such a set-up cost could include, for example, the cost of starting the operation of alternative  $k$  and of shutting down alternative  $j$  when in state  $i$ . Set-up costs can easily be introduced into Models I and II; we illustrate how this is done in Model I for a fixed cost,  $S$ , which is incurred for each change of alternative made. The method is easily generalized to the case in which  $S$  is a function of the state in which the change is made and of the alternatives involved in the change, and is also applicable to the adaptive control model of Chapter 3.

Let  $\underline{\sigma} = (\sigma_1, \dots, \sigma_N)$  denote the policy under which the system is currently operating, where  $\sigma_i$  is the index of the alternative in use in the  $i$ th state ( $\sigma_1 = i, \dots, K_1$ ). We now define the generalised state of the system as  $(i, \Psi, \underline{\sigma})$ , where  $i$  is the physical state of the system ( $i=1, \dots, N$ ),  $\Psi$  indexes the prior distribution of  $\tilde{P}$  ( $\Psi \in \Psi$ ), and  $\underline{\sigma}$  is the policy currently in use ( $\underline{\sigma} \in \Sigma$ ). Let  $v_i(\Psi, \underline{\sigma})$  be the supremum of the expected total discounted reward over an infinite period if the system starts in the generalised state,  $(i, \Psi, \underline{\sigma})$ . The prior distribution function of  $\tilde{P}$  is assumed to belong to a family of distributions closed under consecutive sampling.

If the system is in state  $(i, \Psi, \underline{\sigma})$  and it is decided to sample alternative  $k$ , the supremum of the expected reward is

$$(\delta_{K\sigma_i} - 1)S + \tilde{q}_i^k(\Psi) = \beta c_i^k(\Psi) + \beta \sum_{j=1}^N \tilde{p}_{ij}^k(\Psi) v_j(i_{ij}^k(\Psi), \underline{\sigma}^*(k)) \quad (5.6.1)$$

where  $\underline{\sigma}^*(k) = (\sigma_1^*, \dots, \sigma_N^*)$ , defined by



$$\begin{aligned} \underline{\sigma}_a^* &= \underline{\sigma}_a & a \neq i \\ &= k, & a = i \end{aligned} \quad (5.6.2)$$

is the new policy vector when alternative  $k$  is chosen in state  $i$ ;  $\delta_{ij}$  is the Kronecker delta. The quantities  $\bar{q}_j^k(\gamma)$  and  $\bar{c}_j^k(\gamma)$  are defined by equations (3.1.4) and (5.1.3), respectively.

If it is decided to cease sampling when the system is in state  $(1, \gamma, \underline{\sigma})$  and operate indefinitely under the policy  $\underline{\sigma}' = (\sigma'_1, \dots, \sigma'_N)$ , the expected reward is

$$s \sum_{j=1}^N (\delta_{kj} v_j - 1) + \tilde{V}_1(\underline{\sigma}', \gamma), \quad (5.6.3)$$

where  $\tilde{V}_1(\underline{\sigma}', \gamma)$  is the prior expected discounted reward for operating the system over an infinite period under the policy  $\underline{\sigma}'$ , starting from state 1, when  $P(\underline{P}|\gamma)$  is the prior distribution of  $\underline{P}$ .

Thus, in Model I with a fixed set-up cost,  $v_1(\gamma, \underline{\sigma})$  must satisfy the following functional equation,

$$\begin{aligned} v_1(\gamma, \underline{\sigma}) &= \max \left[ \max_{1 \leq k \leq K_1} \left\{ (\delta_{k \sigma_1} - 1) + \bar{q}_1^k(\gamma) - \bar{c}_1^k(\gamma) + \beta \sum_{j=1}^N \bar{p}_{1j}^k(\gamma) v_j(\bar{T}_{1j}^k(\gamma), \underline{\sigma}^*(k)) \right\} \right. \\ &\quad \left. \max_{\underline{\sigma} \in \Sigma} \left\{ s \sum_{j=1}^N (\delta_{kj} \sigma'_j - 1) + \tilde{V}_1(\underline{\sigma}', \gamma) \right\} \right] \end{aligned} \quad (5.6.4)$$

$1 \leq i, \dots, N$   
 $\gamma \in \Gamma$   
 $\underline{\sigma} \in \Sigma$



## CHAPTER 6

### DISTRIBUTION THEORY

In this chapter we introduce some probability mass functions and density functions which will be required for the next chapter, where we do the prior-posterior and preposterior analysis of a Markov chain observed under the consecutive sampling rule. The Whittle, Whittle-1, and Whittle-2 probability mass functions are defined in Section 6.1 and formulas for their moments are derived. The multivariate beta density function is considered in Section 6.2 and is used to define the matrix beta density function in Section 6.3. Some extensions of the matrix beta distribution are considered in Section 6.4 and the chapter concludes with a discussion of the beta-Whittle probability mass function.

The multivariate beta density function, as defined by equation (6.2.1) below, was introduced by Mardia [30] in 1959; Mosimann [31] has studied the main properties of this distribution. The matrix beta distribution was used by Silver [38], but not under that name. The Whittle and beta-Whittle distributions are original with the present work.

#### 6.1 The Whittle Distribution and Related Distributions.

Let  $x_n = (x_0, x_1, \dots, x_n)$  be the sequence of consecutive observations of the states of a Markov chain over a period of  $n$  transitions, where  $x_0 = u$  is the state of the system prior to the first transition. The range set of the random variables  $\tilde{x}_j$  is the set of integers which index the states of the chain,  $\{1, \dots, N\}$ . It is assumed that the transitions are governed by a known  $N \times N$  stochastic matrix,  $P = [p_{ij}]$ , and that the



distribution of the initial state,  $\tilde{u}$ , is a known stochastic row vector,  
 $p = (p_1, \dots, p_N)$ .

Given a sample outcome,  $x_n$ , we define the statistic  $f_{ij}$  as the number  
of indices  $n \in \{0, 1, \dots, n-1\}$ , such that  $x_n = i$  and  $x_{n+1} = j$  ( $i, j = 1, \dots, N$ ).  
In other words,  $f_{ij}$  is the number of occurrences of a transition from state  
*i* to state *j* in the sample  $x_n$ . Let  $\tilde{F} = [f_{ij}]$ , an  $N \times N$  matrix, be the  
transition count of the sample. Prior to the observation of  $x_n$ ,  $\tilde{F}$  and  
 $\tilde{u}$  are random quantities whose joint distribution is studied in this section.

Let

$$f_{i \cdot} = \sum_{j=1}^N f_{ij} \quad i=1, \dots, N \quad (6.1.1)$$

and

$$f_{\cdot j} = \sum_{i=1}^N f_{ij} \quad j=1, \dots, N \quad (6.1.2)$$

be the row and column sums of  $\tilde{F}$ . With the exception of the initial and  
final transitions, every transition into state *i* in the sample  $x_n$  must be  
followed by a transition out of state *i*. Therefore, the elements of  $\tilde{F}$   
are constrained by the equations

$$f_{i \cdot} - f_{\cdot i} = \delta_{iu} - \delta_{iv}, \quad i=1, \dots, N \quad (6.1.3)$$

where  $u = x_0$  in the initial state and  $v = x_n$  is the final state.

The following lemma shows that, given a transition count  $\tilde{F}$  and an  
initial state  $u$ , the final state of the sample is uniquely determined.  
A similar derivation shows that, given  $\tilde{F}$  and  $v$ ,  $u$  is uniquely determined.  
Lemma 6.1.5, below, shows that  $\tilde{F}$  does not necessarily uniquely determine  
both  $u$  and  $v$ .

Lemma 6.1.1 Let  $n \in \{1, \dots, N\}$  be fixed. If  $\tilde{F}$  is an  $N \times N$  matrix  
of non-negative integers which satisfies the equations

$$f_{i \cdot} - f_{\cdot i} = \delta_{iu} - \delta_{iv} \quad i=1, \dots, N \quad (6.1.4)$$



for some integer  $u \in \{1, \dots, N\}$ , then  $v$  is the only positive integer for which (6.1.4) is true.

**Proof.** The proof is by contradiction. Assume that  $v$  and  $w$  both satisfy (6.1.4). If  $u \neq v$ , then

$$f_{v.} - f_{.v} = \delta_{vu} - \delta_{vv} = -1$$

and, also,

$$f_{v.} - f_{.v} = \delta_{vu} - \delta_{vw} = -\delta_{vw}$$

which implies that  $v = w$ . If  $u = v$  then

$$f_{i.} - f_{.i} = \delta_{iu} - \delta_{iv} = 0 \quad i=1, \dots, N$$

and

$$f_{i.} - f_{.i} = \delta_{iu} - \delta_{iw} \quad i=1, \dots, N$$

Thus,

$$\delta_{iu} = \delta_{iw}, \quad i=1, \dots, N$$

and  $w = u = v$ . Q.E.D.

Let  $I = \{0, 1, 2, \dots\}$  denote the set of all non-negative integers. For fixed  $u \in \{1, \dots, N\}$ ,  $v \in \{1, \dots, N\}$ ,  $n \in \{1, 2, 3, \dots\}$ , and  $p \in \mathcal{P}_N$ , define the following set of  $N \times N$  matrices,  $F = [f_{ij}]$ ,

$$\Phi_N(u, v, n, p) =$$

$$\left\{ F \mid f_{ij} \in I, \sum_{i=1}^N \sum_{j=1}^N f_{ij} = n, f_{i.} - f_{.i} = \delta_{iu} - \delta_{iv}, f_{ij} = 0 \text{ if } p_{ij} = 0, (i, j=1, \dots, N) \right\} \quad (6.1.5)$$

Let

$$\Phi_N(u, n, p) = \bigcup_{v=1}^N \Phi_N(u, v, n, p). \quad u=1, \dots, N \quad (6.1.6)$$

$v=1, 2, \dots$

$p \in \mathcal{P}_N$

It is clear that  $\Phi_N(u, n, p)$  is the set of all possible transition counts  $F$  which can arise from a sample of  $n$  consecutive transitions in a Markov



chain with transition matrix  $P$  and initial state  $u$ .

**6.1.1 The Whittle Distribution.** The  $N \times N$  random matrix  $\tilde{P} = [\tilde{f}_{ij}]$  with range set  $\Phi_N(u, n, P)$  is said to have the Whittle distribution with parameter  $(u, n, P)$  if  $\tilde{P}$  has the joint probability mass function

$$f_w^{(N)}(\tilde{P} | u, n, P) = F_{vu}^* \frac{\prod_{i=1}^N f_{i.1}}{\prod_{i=1}^N \prod_{j=1}^N f_{ij}!} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}} \quad g_e \phi_N(u, n, P)$$

$$= 0, \quad \text{otherwise} \quad (6.1.7)$$

where

$$\begin{aligned} u &= 1, \dots, N \\ n &= 1, 2, 3, \dots \\ P &\in \mathcal{P}_N \end{aligned}$$

The index  $v$  is the unique solution of the equations

$$f_{i.} - f_{.i} = \delta_{iu} - \delta_{iv} \quad i=1, \dots, N$$

and  $F_{vu}^*$  is the  $(v, u)$ th cofactor of the  $N \times N$  matrix  $\tilde{P}^* = [\tilde{f}_{ij}^*]$  defined by

$$\tilde{f}_{ij}^* = \delta_{ij} - \frac{f_{ij}}{f_{i.}} \quad f_{i.} > 0 \quad (6.1.8a)$$

$$= \delta_{ij}. \quad f_{i.} = 0 \quad (6.1.8b)$$

Since, in (6.1.7), there may be some  $p_{ij} = 0$ , we use the convention  $0^0 = 1$ .

We have called the mass function (6.1.7) the Whittle distribution because Whittle [41] was the first to show that

$$\sum_{\substack{u \\ \in \\ \Phi_N(u, n, P)}} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}} = \frac{\prod_{i=1}^N \prod_{j=1}^N f_{ij}!}{F_{vu}^* \prod_{i=1}^N f_{i.}!} \quad (6.1.9)$$

Whittle's derivation of (6.1.9), and subsequent proofs of this relation by Dawson and Good [15] and by Goodman [21], were obtained under the



restriction  $f_{1i} > 0$  ( $i=1, \dots, N$ ). Billingsley [10], in a particularly elegant proof of (6.1.9), did not require this restriction.

**6.1.2 Moments of the Whittle Distribution.** We now derive expressions for the means, variances, and covariances of the elements of  $\tilde{P}$ . Before presenting these results, however, it is necessary to summarize certain facts from the theory of matrices.

Let  $P$  be an  $N \times N$  matrix with eigenvalues  $\lambda_1, \dots, \lambda_N$ , assumed to be distinct. Let  $g(x)$  be an arbitrary scalar polynomial,  $a_0 + a_1x + \dots + a_Nx^N$ , and let  $g(P)$  be the corresponding matrix polynomial,  $a_0I + a_1P + \dots + a_NP^N$ . Sylvester's Theorem states that

$$g(P) = \sum_{k=1}^N g(\lambda_k) A^{(k)}, \quad (6.1.10)$$

where the  $N \times N$  matrices  $A^{(k)}$  are defined by the expression

$$A^{(k)} = [a_{ij}^{(k)}] = \frac{\prod_{j \neq k} (\lambda_j I - P)}{\prod_{j \neq k} (\lambda_j - \lambda_k)}. \quad (6.1.11)$$

$k=1, \dots, N$

These matrices have the following properties:

$$A^{(i)} A^{(j)} = 0, \quad i \neq j \quad (6.1.12a)$$

$$[A^{(i)}]^2 = A^{(i)}. \quad i=1, \dots, N \quad (6.1.12b)$$

A formula similar to (6.1.10), called the confluent form of Sylvester's theorem, is available in the case of repeated eigenvalues.

If  $P$  is an ergodic stochastic matrix--i.e., if  $P$  is the transition matrix of a single non-periodic Markov chain--then exactly one eigenvalue has the value unity and all other eigenvalues have modulus less than unity. We shall adopt the convention that  $\lambda_1$  is the unit root:



$$\lambda_1 = 1 \quad (6.1.13a)$$

$$|\lambda_i| < 1, \quad i=2, \dots, N \quad (6.1.13b)$$

Then the matrix  $\underline{A}^{(1)}$  is an  $N \times N$  matrix each row of which is the steady state vector  $\underline{\pi} = (\pi_1, \dots, \pi_N)$  defined by the relation  $\underline{\pi} = \underline{\pi}P$ .

**Theorem 6.1.2** If the  $N \times N$  random matrix  $\tilde{F}$  has the Whittle distribution with parameter  $(u, n, P)$ , then the expected value of  $\tilde{f}_{ij}$  is

$$E[\tilde{f}_{ij}] = \tilde{f}_{ij}(u, n) = \sum_{k=0}^{n-1} p_{ui}^{(k)} p_{ij}, \quad \begin{matrix} i, j = 1, \dots, N \\ n = 1, 2, 3, \dots \\ u = 1, \dots, N \end{matrix} \quad (6.1.14)$$

where  $p_{ui}^{(k)}$  is the  $(u, i)$ th element of  $P^k$ . If, furthermore,  $P$  is ergodic and the eigenvalues,  $\lambda_1, \dots, \lambda_N$ , of  $P$  are distinct, then the expected value of  $f_{ij}$  has the spectral representation

$$\tilde{f}_{ij}(u, n) = p_{ij} [n\pi_1 + \sum_{m=2}^N \frac{i-\lambda_m^n}{i-\lambda_m} a_{ui}^{(m)}], \quad \begin{matrix} i, j = 1, \dots, N \\ n = 1, 2, 3, \dots \\ u = 1, \dots, N \end{matrix} \quad (6.1.15)$$

where  $\underline{a}^{(m)} = [a_{1j}^{(m)}, \dots, a_{Nj}^{(m)}]$  is defined by (6.1.11).

**Proof.** Let  $f_{ij}(u, n)$  be the number of transitions from  $i$  to  $j$  in a sample of  $n$  transitions which has initial state  $u$ . Prior to the observation of the sample,  $\tilde{f}_{ij}(u, n)$  is a random variable. If the system starts in state  $u$  and the first transition is to state  $k$ , then  $\tilde{f}_{ij}(u, n)$  satisfies the equations

$$\tilde{f}_{ij}(u, n) = \delta_{ui} \delta_{kj} + \tilde{f}_{ij}(k, n-1), \quad \begin{matrix} m = 2, 3, \dots \\ k = 1, \dots, N \end{matrix} \quad (6.1.16a)$$

and

$$\tilde{f}_{ij}(u, 1) = \delta_{ui} \delta_{kj}. \quad k = 1, \dots, N \quad (6.1.16b)$$

Thus,  $\tilde{f}_{ij}(u, n)$  satisfies the equations



$$\tilde{f}_{ij}(u, n) = p_{uj} \delta_{ui} + \sum_{k=1}^N p_{uk} \tilde{f}_{kj}(k, n-1) \quad n=2, 3, \dots \quad (6.1.17a)$$

$$\tilde{f}_{ij}(u, 1) = p_{uj} \delta_{ui}. \quad (6.1.17b)$$

We shall prove inductively that

$$\tilde{f}_{ij}(u, n) = \sum_{k=0}^{n-1} p_{ui}^{(k)} p_{kj}. \quad \begin{matrix} i, j = 1, \dots, N \\ n = 1, 2, 3, \dots \\ u = 1, \dots, N \end{matrix} \quad (6.1.18)$$

Since  $\delta_{ui} p_{ij} = \delta_{ui} p_{uj}$ , (6.1.17) is satisfied by (6.1.18) for  $n = 1$ . Assume (6.1.18) holds for  $n$ . Then, using (6.1.17a),

$$\begin{aligned} \tilde{f}_{ij}(u, n+1) &= p_{uj} \delta_{ui} + \sum_{k=1}^N p_{uk} \sum_{m=0}^{n-1} p_{im}^{(m)} p_{kj} \\ &= p_{uj} \delta_{ui} + \sum_{m=0}^{n-1} p_{ui}^{(m+1)} p_{kj} \\ &= \sum_{k=0}^n p_{ui}^{(k)} p_{kj}. \end{aligned} \quad (6.1.19)$$

proving the induction.

If all the eigenvalues of  $P$  are distinct, Sylvester's Theorem yields

$$p_{ui}^{(k)} = \sum_{m=1}^N \lambda_m^k a_{ui}^{(m)}. \quad k=0, 1, 2, \dots \quad (6.1.20)$$

If, furthermore,  $P$  is ergodic and  $\lambda_1 = 1$  is the only eigenvalue of unit modulus, equation (6.1.18) can be written

$$\begin{aligned} \tilde{f}_{ij}(u, n) &= \sum_{k=0}^{n-1} \sum_{m=1}^N \lambda_m^k a_{ui}^{(m)} p_{kj} \\ &= p_{ij} [n \pi_2 + \sum_{m=2}^N \frac{1 - \lambda_m^n}{1 - \lambda_m} a_{ui}^{(m)}]. \end{aligned} \quad (6.1.21)$$

$i, j = 1, \dots, N$   
 $m = 1, 2, 3, \dots$   
 $u = 1, \dots, N$

Q.E.D.



Theorem 6.1.3 If the  $N \times N$  random matrix  $\tilde{F}$  has the Whittle distribution with parameter  $(u, n, P)$ , then, for  $a, \beta, \gamma, \delta = 1, \dots, N$ , the covariance between  $\tilde{F}_{ab}$  and  $\tilde{F}_{\gamma\delta}$  is,

$$\text{cov} [\tilde{F}_{ab}, \tilde{F}_{\gamma\delta}] =$$

$$\tilde{F}_{ab}(u, n) [\delta_{ay} \delta_{\beta\delta} - \tilde{F}_{\gamma\delta}(u, n)] + \sum_{k=1}^{n-1} [p_{uz}^{(n-1-k)} p_{ab} \tilde{F}_{\gamma\delta}(\beta, k) + p_{uy}^{(n-1-k)} p_{\gamma\delta} \tilde{F}_{ab}(\delta, k)]$$

$n=2, 3, \dots \quad (6.1.22a)$

$$= p_{ab} p_{\gamma\delta} [\delta_{ay} \delta_{\beta\delta} - \delta_{\gamma\delta} p_{ab}] \quad n=1 \quad (6.1.22b)$$

If  $P$  is ergodic and the eigenvalues of  $P$ ,  $\lambda_1, \dots, \lambda_N$ , are all distinct, the covariance of  $\tilde{F}_{ab}$  and  $\tilde{F}_{\gamma\delta}$  has the spectral representation

$$\begin{aligned} \text{cov} [\tilde{F}_{ab}, \tilde{F}_{\gamma\delta}] &= p_{ab} p_{\gamma\delta} [n \pi_a + \sum_{m=2}^N \frac{1-\lambda_m^n}{1-\lambda_m} a_{um}^{(n)}] \\ &- p_{ab} p_{\gamma\delta} \left\{ n \sum_{m=2}^N \left[ \frac{1-\lambda_m^n}{1-\lambda_m} (\pi_a a_{uy}^{(n)} + \pi_y a_{ux}^{(n)}) \right. \right. \\ &\quad \left. \left. + \sum_{m=2}^N \sum_{j=2}^N \left( \frac{1-\lambda_m^n}{1-\lambda_m} \right) \left( \frac{1-\lambda_j^n}{1-\lambda_j} \right) a_{uj}^{(n)} a_{uy}^{(j)} \right] \right\} \\ &+ p_{ab} p_{\gamma\delta} \left\{ -n \pi_a \pi_y + \sum_{m=2}^N \frac{n-1-n\lambda_m+\lambda_m^n}{(1-\lambda_m)^2} [\pi_a (a_{uy}^{(n)} + a_{\beta y}^{(n)}) \right. \\ &\quad \left. + \pi_y (a_{ua}^{(n)} + a_{\delta u}^{(n)})] + \sum_{m=2}^N \frac{1-n\lambda_m^{n-1}+(n-1)\lambda_m^n}{(1-\lambda_m)^2} (a_{ua}^{(n)} a_{\beta y}^{(n)} + a_{wy}^{(n)} a_{\delta u}^{(n)}) \right. \\ &\quad \left. + \sum_{m=2}^N \sum_{\substack{j=1 \\ j \neq m}}^N \frac{a_{uj}^{(n)} a_{\beta y}^{(j)} + a_{uy}^{(n)} a_{\delta u}^{(j)}}{1-\lambda_j} \left[ \frac{1-\lambda_m^{n-1}}{1-\lambda_m} - \frac{\lambda_1(\lambda_m^{n-1}-\lambda_1^{n-1})}{\lambda_m-\lambda_1} \right] \right\} \quad (6.1.23) \end{aligned}$$

when  $n > 1$ .

**Proof.** If  $\tilde{F}_{ab}(u, n)$  is the number of transitions from state  $u$  to state  $\beta$  in a sample of  $n$  transitions when the chain starts in state  $u$ , equation (6.1.16) and the relation  $\delta_{ik} \delta_{kj} = \delta_{ij} \delta_{ki}$  imply that, if the



first transition is from  $u$  to  $k$ ,

$$\tilde{F}_{\alpha\beta}(u,n) \tilde{F}_{\gamma\delta}(u,n) =$$

$$\delta_{\alpha\gamma} \delta_{\beta\delta} \delta_{\alpha u} \delta_{\beta k} + \delta_{uv} \delta_{k\delta} \tilde{F}_{\alpha\beta}(k,n-1) + \delta_{ua} \delta_{k\beta} \tilde{F}_{\gamma\delta}(k,n-1) + \tilde{F}_{\alpha\beta}(k,n-1) \tilde{F}_{\gamma\delta}(k,n-1),$$

$$m=2, 3, \dots \quad (6.1.24a)$$

$$k=1, \dots, N$$

$$= \delta_{\alpha\gamma} \delta_{\beta\delta} \delta_{\alpha u} \delta_{\beta k}. \quad m=1 \quad (6.1.24b)$$

$$k=1, \dots, N$$

Let

$$\sigma_{\alpha\beta\gamma\delta}(u,n) = E[\tilde{F}_{\alpha\beta}(u,n) \tilde{F}_{\gamma\delta}(u,n)]. \quad (6.1.25)$$

Then, using equation (6.1.24), it is seen that  $\sigma_{\alpha\beta\gamma\delta}(u,n)$  satisfies the equations

$$\Gamma_{\alpha\beta\gamma\delta}(u,n) =$$

$$\delta_{\alpha\gamma} \delta_{\beta\delta} \delta_{\alpha u} p_{u\beta} + \delta_{uv} p_{u\delta} \tilde{F}_{\alpha\beta}(s,n-1) + \delta_{us} p_{u\beta} \tilde{F}_{\gamma\delta}(s,n-1) + \sum_{k=1}^N p_{uk} \sigma_{\alpha\beta\gamma\delta}(k,n-1),$$

$$n=2, 3, \dots \quad (6.1.26a)$$

$$= \delta_{\alpha\gamma} \delta_{\beta\delta} \delta_{\alpha u} p_{u\beta}. \quad n=1 \quad (6.1.26b)$$

We shall show that

$$\sigma_{\alpha\beta\gamma\delta}(u,n) =$$

$$\delta_{\alpha\gamma} \delta_{\beta\delta} \tilde{F}_{\alpha\beta}(u,n) + \sum_{k=1}^{n-1} [p_{uu}^{(n-1-k)} p_{\alpha\beta} \tilde{F}_{\gamma\delta}(\beta,k) + p_{uv}^{(n-1-k)} p_{\gamma\delta} \tilde{F}_{\alpha\beta}(\delta,k)],$$

$$u=1, \dots, N \quad (6.1.27a)$$

$$v=2, 3, \dots$$

$$= \delta_{\alpha\gamma} \delta_{\beta\delta} \tilde{F}_{\alpha\beta}(u,1). \quad u=1, \dots, N \quad (6.1.27b)$$

$$v=1$$

In which case equation (6.1.22) follows.



It is clear from (6.1.17b) that (6.1.27b) equals (6.1.26b). The case  $n=2, 3, \dots$  will be proven by induction. For  $n=2$ , it is easily verified that (6.1.26a) is equal to the expression in (6.1.27a). Assume (6.1.27a) satisfies (6.1.26a) for  $n$ . Then, using (6.1.14),

$$\begin{aligned}
 \sigma_{\alpha\beta\gamma\delta}(u, n+1) &= \delta_{\alpha\gamma}\delta_{\beta\delta}\delta_{\alpha u}p_{u\beta} + \delta_{uy}p_{ub}\tilde{F}_{ab}(b, n) + \delta_{ua}p_{ub}\tilde{F}_{yb}(b, n) \\
 &\quad + \sum_{k=1}^N p_{uk} \left[ \delta_{\alpha\gamma}\delta_{\beta\delta} \sum_{m=0}^{n-1} p_{ka}^{(m)} p_{ab} + \sum_{m=1}^{n-1} [p_{ka}^{(n-1-m)} p_{ab}\tilde{F}_{yb}(b, m) + p_{ky}^{(n-1-m)} p_{yb}\tilde{F}_{ab}(b, m)] \right] \\
 &= \delta_{\alpha\gamma}\delta_{\beta\delta}\delta_{\alpha u}p_{u\beta} + \delta_{uy}p_{ub}\tilde{F}_{ab}(b, n) + \delta_{ua}p_{ub}\tilde{F}_{yb}(b, n) \\
 &\quad + \delta_{\alpha\gamma}\delta_{\beta\delta} \sum_{m=0}^{n-1} p_{ua}^{(m+1)} p_{ab} + \sum_{m=1}^{n-1} [p_{ua}^{(n-m)} p_{ab}\tilde{F}_{yb}(b, m) + p_{uy}^{(n-m)} p_{yb}\tilde{F}_{ab}(b, m)] \\
 &= \delta_{\alpha\gamma}\delta_{\beta\delta}\tilde{F}_{ab}(u, n+1) + \sum_{k=1}^n [p_{ua}^{(n-k)} p_{ab}\tilde{F}_{yb}(b, k) + p_{uy}^{(n-k)} p_{yb}\tilde{F}_{ab}(b, k)], \tag{6.1.28}
 \end{aligned}$$

proving the induction.

If  $\underline{P}$  is ergodic and has distinct eigenvalues, equations (6.1.15) and (6.1.20) for the spectral representations of  $\tilde{F}_{1j}(u, n)$  and  $p_{ui}^{(k)}$  can be used in (6.1.22a) to obtain, for  $n=2, 3, \dots$

$$\begin{aligned}
 \text{cov}[\tilde{F}_{ab}, \tilde{F}_{yb}] &= p_{ab}[n\pi_a + \sum_{m=2}^N \frac{1-\lambda_m^n}{1-\lambda_m} a_{ua}^{(m)}](\delta_{\alpha\gamma}\delta_{\beta\delta} - p_{yb}[n\pi_y + \sum_{m=2}^N \frac{1-\lambda_m^n}{1-\lambda_m} a_{uy}^{(m)}]) \\
 &\quad + p_{ab}p_{yb} \sum_{k=1}^{n-1} \sum_{m=1}^N \lambda_m^{(n-1-k)} [a_{ua}^{(m)}(k\pi_y + \sum_{j=2}^N \frac{1-\lambda_j^k}{1-\lambda_j} a_{by}^{(j)}) + a_{uy}^{(m)}(k\pi_a + \sum_{j=2}^N \frac{1-\lambda_j^k}{1-\lambda_j} a_{ba}^{(j)})]. \tag{6.1.29}
 \end{aligned}$$

Multiplying out equation (6.1.29) and using the relations

$$\sum_{k=1}^{n-1} \frac{k\lambda_m^{(n-1-k)}}{(1-\lambda_m)^2} = \frac{n-1-n\lambda_m + \lambda_m^n}{(1-\lambda_m)^2}, \quad m=2, \dots, N \tag{6.1.30}$$



$$\sum_{k=1}^{n-1} (1-\lambda_j^k) = \frac{\lambda_j^{n-1} + \lambda_j^n}{1-\lambda_j}, \quad j=2, \dots, N \quad (6.1.31)$$

$$\sum_{k=1}^{m-1} \lambda_m^{(n-1-k)} (1-\lambda_j^k) = \frac{1-\lambda_m^{m-1}}{1-\lambda_m} - \frac{\lambda_j (\lambda_m^{m-1} - \lambda_j^{m-1})}{\lambda_m - \lambda_j}. \quad (6.1.32)$$

$j=2, \dots, N$   
 $j \neq m$

and

$$\sum_{k=1}^{m-1} \lambda_m^{(n-1-k)} (1-\lambda_m^k) = \frac{1-\lambda_m^{m-1} + (n-1)\lambda_m^n}{1-\lambda_m}, \quad m=2, \dots, N \quad (6.1.33)$$

equation (6.1.23) is obtained. Q.E.D.

In Theorems 6.1.2 and 6.1.3 the spectral representations provide an efficient method of computing the means, variances, and covariances of elements of  $\tilde{\Gamma}$ . This method is particularly useful as the parameter  $n$  becomes large and, in fact, leads to relatively simple approximations for  $\tilde{x}_{ij}(u, n)$  and  $\text{cov} [\tilde{f}_{\alpha\beta}, \tilde{f}_{\gamma\delta}]$  when  $n$  is sufficiently large, as is shown in the following corollary.

Corollary 6.1.4 If the  $N \times N$  random matrix  $\tilde{\Gamma}$  has the Whittle distribution with parameter  $(u, n, P)$ , where  $P$  is ergodic and has distinct eigenvalues, then, for large  $n$ , the expected value of  $\tilde{f}_{ij}$  and the covariance between  $\tilde{f}_{\alpha\beta}$  and  $\tilde{f}_{\gamma\delta}$  are given by the following asymptotic expressions:

$$\tilde{x}_{ij}(u, n) \approx p_{ij} [n\pi_i + \sum_{m=2}^N \frac{a_m(u)}{1-\lambda_m}], \quad i, j=1, \dots, N \quad (6.1.34)$$



$$\text{cov} [\tilde{F}_{\alpha\beta}, \tilde{F}_{\gamma\delta}] \approx$$

$$\begin{aligned}
 & n[\pi_{\alpha\beta\gamma\delta}\pi_{\alpha\gamma\delta\beta} - \pi_{\alpha\beta\gamma\delta}\pi_{\gamma\beta\alpha\delta} + p_{\alpha\beta}p_{\gamma\delta} \sum_{m=2}^N \frac{\pi_{\alpha\beta\gamma}^{(m)} + \pi_{\gamma\delta\alpha}^{(m)}}{1-\lambda_m}] \\
 & + p_{\alpha\beta}p_{\alpha\gamma\delta\beta} \sum_{m=2}^N \frac{a_{\alpha\alpha}^{(m)}}{1-\lambda_m} = p_{\alpha\beta}p_{\gamma\delta} \sum_{m=2}^N \frac{\pi_{\alpha}(a_{\alpha\gamma}^{(m)} + a_{\beta\gamma}^{(m)}) + \pi_{\gamma}(a_{\alpha\alpha}^{(m)} + a_{\delta\alpha}^{(m)})}{(1-\lambda_m)^2} \\
 & + p_{\alpha\beta}p_{\gamma\delta} \sum_{m=2}^N \sum_{j=2}^N \frac{a_{\alpha\alpha}^{(m)}a_{\beta\gamma}^{(j)} + a_{\alpha\gamma}^{(m)}a_{\beta\alpha}^{(j)} - a_{\alpha\alpha}^{(m)}a_{\beta\gamma}^{(j)}}{(1-\lambda_j)(1-\lambda_m)}. \tag{6.1.35}
 \end{aligned}$$

$$\alpha, \beta, \gamma, \delta = 1, \dots, N$$

Proof. Equations (6.1.34) and (6.1.35) are obtained by letting  $n$  become large in (6.1.15) and (6.1.23), dropping terms of order  $\lambda_m^n$  ( $m=2, \dots, N$ ), and noting that

$$\lim_{n \rightarrow \infty} n\lambda_m^n = 0. \quad m=2, \dots, N \tag{6.1.36}$$

Q.E.D.

**6.1.3 The Whittle-i Distribution.** Let  $\tilde{u}$  be a random integer with range set  $\{1, \dots, N\}$  and let  $\tilde{F} = [\tilde{F}_{ij}]$  be an  $N \times N$  random matrix with range set  $\Phi_N(u, n, p)$ . The ordered pair  $(\tilde{u}, \tilde{F})$  is said to have the Whittle-i distribution with parameter  $(p, n, \underline{p})$  if  $(\tilde{u}, \tilde{F})$  has the joint probability mass function

$$\begin{aligned}
 f_{W1}^{(N)}(u, \underline{p} \mid p, n, \underline{p}) &= p_u f_W^{(N)}(\underline{p} \mid u, n, \underline{p}) & u=1, \dots, N \\
 & \frac{p_u}{\underline{p}} \Phi_N(u, n, \underline{p}) & \underline{p} \in \Phi_N(u, n, \underline{p}) \\
 & = 0, & \text{otherwise} \tag{6.1.37}
 \end{aligned}$$

where  $\underline{p} = (p_1, \dots, p_N)$  is a stochastic row vector,  $n=1, 2, 3, \dots$ , and  $p \in \mathcal{P}_N$ .



Since  $p_u \geq 0$  and  $\sum_{u=1}^N p_u = 1$ , it is clear that

$$f_{W1}^{(N)}(u, F \mid \underline{p}, \underline{n}, \underline{P}) \geq 0$$

and, using equation (6.1.9),

$$\sum_{u=1}^N \sum_{\substack{F \in \phi_N \\ u}} f_{W1}^{(N)}(u, F \mid \underline{p}, \underline{n}, \underline{P}) = 1. \quad (6.1.38)$$

It is readily seen that, if  $(\tilde{u}, \tilde{F})$  has the Whittle-1 distribution with parameter  $(\underline{p}, \underline{n}, \underline{P})$ , the marginal distribution of  $\tilde{u}$  is

$$\begin{aligned} P[u \mid \underline{p}] &= p_u, & u=1, \dots, N \\ &= 0. & \text{otherwise} \end{aligned} \quad (6.1.39)$$

The marginal distribution of  $\tilde{F}$  is considered in the remaining paragraphs of this section.

**6.1.4 The Whittle-2 Distribution.** Let  $\tilde{F}$  be an  $N \times N$  random matrix with range set

$$\Phi_N^{*}(\underline{n}, \underline{P}) = \bigcup_{\substack{u=1 \\ F \in \mathcal{S}_N}}^N \Phi_N^{(N)}(u, F \mid \underline{p}, \underline{n}, \underline{P}). \quad n=1, 2, 3, \dots \quad (6.1.40)$$

The Whittle-2 distribution with parameter  $(\underline{p}, \underline{n}, \underline{P})$  is defined as the marginal distribution of  $\tilde{F}$  when  $(\tilde{u}, \tilde{F})$  has the Whittle-1 distribution with parameter  $(\underline{p}, \underline{n}, \underline{P})$ ;

$$\begin{aligned} f_{W2}^{(N)}(F \mid \underline{p}, \underline{n}, \underline{P}) &= \sum_{u=1}^N f_{W1}^{(N)}(u, F \mid \underline{p}, \underline{n}, \underline{P}) \cdot \underline{\underline{p}} \cdot \Phi_N^{*}(\underline{n}, \underline{P}) \\ &= 0, & \text{otherwise} \end{aligned} \quad (6.1.41)$$

where  $\underline{p}$  is an  $N$ -dimensional stochastic row vector,  $n=1, 2, \dots$ , and  $\underline{\underline{p}} \in \mathcal{S}_N$ .

It is clear from the definition (6.1.41) and the fact that

$f_{W1}^{(N)}(u, F \mid \underline{p}, \underline{n}, \underline{P})$  is a probability mass function that

$$f_{W2}^{(N)}(F \mid \underline{p}, \underline{n}, \underline{P}) \geq 0$$



and

$$\sum_{\substack{F \in \Phi_N^*(n, P) \\ F \neq \phi}} \mathbb{E}_{W2}^{(N)}(F | p_0, n, P) = 1.$$

Before deriving an explicit formula for  $\mathbb{E}_{W2}^{(N)}(F | p_0, n, P)$ , a preliminary lemma is required. To this end let  $\Phi_N^*(n, P)$  be partitioned into two sets,  $\Phi_{N1}^*(n, P)$  and  $\Phi_{N2}^*(n, P)$ , defined as

$$\Phi_{N1}^*(n, P) = \left\{ F \mid F \in \Phi_N^*(n, P), f_{i,i} = f_{.i,.i} \text{ } (i=1, \dots, N) \right\}, \quad (6.1.42)$$

$$\Phi_{N2}^*(n, P) = \Phi_N^*(n, P) - \Phi_{N1}^*(n, P). \quad (6.1.43)$$

$\Phi_{N1}^*(n, P)$  is the set of all transition counts which start and end in the same state and  $\Phi_{N2}^*(n, P)$  is the set of all other transition counts in  $\Phi_N^*(n, P)$ . Both sets are non-empty.

**Lemma 6.1.4** Let the sets  $\Phi_{N1}^*(n, P)$  and  $\Phi_{N2}^*(n, P)$  be defined by equations (6.1.42) and (6.1.43). If  $F \in \Phi_{N1}^*(n, P)$  there are exactly  $N$  pairs of integers,  $(x, y) = (u, u)$ ,  $u=1, \dots, N$ , which satisfy the equations

$$f_{i,i} - f_{.i,.i} = \delta_{ix} - \delta_{iy}. \quad i=1, \dots, N \quad (6.1.44)$$

If, on the other hand,  $F \in \Phi_{N2}^*(n, P)$  there is a unique solution,  $(x, y) = (u, v)$ , where  $u \neq v$ , to (6.1.44).

**Proof.** If  $F \in \Phi_{N1}^*(n, P)$ , then

$$f_{i,i} - f_{.i,.i} = 0, \quad i=1, \dots, N$$

and (6.1.44) becomes

$$\delta_{ix} - \delta_{iy} = 0. \quad i=1, \dots, N$$

These equations are satisfied by  $x = y = u$  ( $u=1, \dots, N$ ) and are not satisfied by any pair  $(x, y)$  such that  $x \neq y$ .



If  $\frac{f}{\phi} \in \phi_{N_2}^*(n, p)$  there is, by the definition of  $\phi_{N_2}^*(n, p)$ , at least one solution,  $(u, v)$ , to (6.1.44) with  $u \neq v$ . Assume  $(u^*, v^*)$  also satisfies (6.1.44). If  $u^* = u$ , Lemma 6.1.1 implies  $v^* = v$ . Assume  $u^* \neq u$ . Then, if  $v^* \neq v$ , substitution of  $(u, v)$  in (6.1.44) yields, for  $i = v$ ,

$$f_{v_i} - f_{u_i} = -1$$

while  $(u^*, v^*)$  substituted into (6.1.44) gives

$$f_{v_i} - f_{u_i} = \delta_{vu}^*$$

a contradiction. If  $v^* = v$ , then  $(u, v)$  substituted into (6.1.44) with  $i = u$  implies

$$f_{u_i} - f_{v_i} = 1$$

and, since  $v^* = v \neq u$ ,  $(u^*, v^*)$  substituted into (6.1.44) yields

$$f_{u_i} - f_{v_i} = \delta_{vu}^*,$$

which contradicts the assumption  $u^* \neq u$ . Thus,  $(u^*, v^*) = (u, v)$ . Q.E.D.

**Theorem 6.1.6** Let  $\tilde{\mathbf{F}}$  be an  $N \times N$  random matrix with range set  $\phi_{N_2}^*(n, p)$  which has the Whittle-2 distribution with parameter  $(p, n, \underline{p})$ . Then the probability mass function of  $\tilde{\mathbf{F}}$  is given by

$$f_{W2}^{(N)}(\tilde{\mathbf{F}} | p, n, \underline{p}) = \left( \sum_{i=1}^N p_i F_{ii} \right)^N \frac{\prod_{i=1}^N \prod_{j=1}^N f_{ij}^*}{\prod_{i=1}^N \prod_{j=1}^N f_{ij}!} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^* f_{ij}, \quad \frac{f}{\phi} \in \phi_{N_2}^*(n, p)$$

$$= p_u F_{vu} \frac{\prod_{i=1}^N \prod_{j=1}^N f_{ij}^*}{\prod_{i=1}^N \prod_{j=1}^N f_{ij}!} \prod_{i=1}^N \prod_{j=1}^N p_{ij}^* f_{ij}, \quad \frac{f}{\phi} \in \phi_{N_2}^*(n, p)$$

$$= 0, \quad \text{otherwise} \quad (6.1.45)$$



where  $\hat{F}_{xy}^*$  is the  $(x,y)$ th cofactor of the matrix  $\hat{F}^*$  defined by equation (6.1.8) and, when  $\hat{P} \in \phi_{N_1}^*(n, P)$ ,  $(u, v)$  is the unique solution to equation (6.1.44).

Proof. By definition,  $\phi_{N_1}^*(n, P)$  and  $\phi_{N_2}^*(n, P)$  are mutually exclusive sets and together exhaust the range set,  $\phi_N^*(n, P)$ . If  $\hat{P} \in \phi_{N_1}^*(n, P)$  then, by Lemma 6.1.5,  $\hat{P} \in \phi_N^*(i, n, P)$ ,  $i=1, \dots, N$  and

$$z_W^{(N)}(\hat{F}^* | i, n, P) > 0, \quad i=1, \dots, N$$

which yields the first line of equation (6.1.45). If  $\hat{P} \in \phi_{N_2}^*(n, P)$ , Lemma 6.1.5 implies there is exactly one value of  $u$  in the range  $\{1, \dots, N\}$  such that

$$z_W^{(N)}(\hat{F}^* | u, n, P) > 0,$$

which yields the second line of (6.1.45). Q.E.D.

**6.1.5 Moments of the Whittle-2 Distribution.** In this paragraph we derive formulas for the expected value of  $\tilde{f}_{ij}$  and for the covariance between  $\tilde{f}_{ij}$  and  $\tilde{f}_{\gamma\delta}$  when  $\hat{F}^*$  has the Whittle-2 distribution with parameter  $(p, n, P)$ . When  $P$  is ergodic and  $p = \frac{1}{2}$ , the steady state distribution corresponding to  $P$ , particularly simple formulas result. Related moments have been derived by other authors. Anderson and Goodman [1], assuming that many Markov chains which are governed by the same transition matrix are simultaneously observed, find expressions for the means, variances, and covariances of  $\tilde{n}_{ij}(t)$ , the number of systems making a transition from state  $i$  to state  $j$  on the  $t$ -th transition ( $i, j=1, \dots, N$ ). Good [20] has derived formulas for the mean vector and variance-covariance matrix of the frequency counts,  $\tilde{\mathbf{Z}} = (\tilde{Z}_1, \dots, \tilde{Z}_N)$ , where



$$\tilde{f}_{1j} = \sum_{j=1}^N \tilde{f}_{1j} \quad i=1, \dots, N$$

is the number of times the system is observed to be in state 1 (including the initial state, but not the final state) in a sample of  $n$  consecutive transitions, assuming that the distribution of the initial state is  $\underline{\pi}_0$ , the steady state distribution corresponding to  $P$ . Our equations (6.1.53) and (6.1.54), when summed over  $j$  and over  $\beta$  and  $\delta$ , respectively, reduce to Good's formulae.

**Theorem 6.1.2** Let the  $N \times N$  random matrix  $\tilde{F}$  have the Whittle-2 distribution with parameter  $(p, n, g)$ . Then the expected value of  $\tilde{f}_{2j}$  is

$$E[\tilde{f}_{2j}] = p_{1j} \sum_{k=0}^{n-1} \sum_{u=1}^N p_u p_{ui}^{(k)}. \quad i, j=1, \dots, N \quad (6.1.46)$$

If  $P$  is ergodic and has distinct eigenvalues,  $\lambda_1, \dots, \lambda_N$ , then  $E[\tilde{f}_{2j}]$  has the spectral representation

$$E[\tilde{f}_{1j}] = n\pi_{1j} + p_{1j} \sum_{m=2}^N \frac{1-\lambda_m^N}{1-\lambda_m} \sum_{u=1}^N p_u a_{uj}^{(m)}, \quad i, j=1, \dots, N \quad (6.1.47)$$

where the  $N \times N$  matrices  $A^{(m)} = [a_{ij}^{(m)}]$  are defined by (6.1.11).

**Proof.** Since

$$E[\tilde{f}_{1j}] = \sum_{u=1}^N p_u \tilde{f}_{1j}(u, n), \quad (6.1.48)$$

equation (6.1.46) follows immediately from equation (6.1.14). When  $P$  is ergodic with distinct eigenvalues, (6.1.47) follows from (6.1.15). Q.E.D.

**Theorem 6.1.3** Let the  $N \times N$  random matrix  $\tilde{F}$  have the Whittle-2 distribution with parameter  $(p, n, g)$ . Then, for  $\alpha, \beta, \gamma, \delta = 1, \dots, N$ , the covariance between  $\tilde{f}_{\alpha\beta}$  and  $\tilde{f}_{\gamma\delta}$  is



$$\text{cov}[\tilde{x}_{\alpha\beta}, \tilde{x}_{\gamma\delta}] = E[\tilde{x}_{\alpha\beta}](\delta_{\alpha\gamma}\delta_{\beta\delta} - E[\tilde{x}_{\gamma\delta}]) + \sum_{k=1}^{n-1} [p_{\alpha\beta} \tilde{x}_{\gamma\delta}(\beta, k) \sum_{u=1}^N p_u p_{u\alpha}^{(n-1-k)}]$$

$$+ p_{\gamma\delta} \tilde{x}_{\alpha\beta}(\beta, k) \sum_{u=1}^N p_u p_{uy}^{(n-1-k)}], \quad (6.1.49a)$$

$n=2, 3, \dots$

$$= E[\tilde{x}_{\alpha\beta}](\delta_{\alpha\gamma}\delta_{\beta\delta} - E[\tilde{x}_{\gamma\delta}]). \quad n=1 \quad (6.1.49b)$$

If  $P$  is ergodic with distinct eigenvalues,  $\lambda_1, \dots, \lambda_N$ , and if

$$\underline{b}_m^{(n)} = p_A^{(n)} = (b_1^{(n)}, \dots, b_N^{(n)}), \quad n=2, \dots, N \quad (6.1.50)$$

where  $\underline{b}_m^{(n)}$  is defined by equation (6.1.11), then (6.1.49a) has the spectral representation

$$\text{cov}[\tilde{x}_{\alpha\beta}, \tilde{x}_{\gamma\delta}] =$$

$$\begin{aligned} & p_{\alpha\beta} \delta_{\alpha\gamma} \delta_{\beta\delta} [n\pi_\alpha + \sum_{m=2}^N \frac{1-\lambda_m^n}{1-\lambda_m} b_m^{(n)}] - p_{\alpha\beta} p_{\gamma\delta} \left\{ n \sum_{m=2}^N \left[ \frac{1-\lambda_m^n}{1-\lambda_m} (\pi_\alpha b_\gamma^{(n)} + \pi_\gamma b_\alpha^{(n)}) \right] \right. \\ & \left. + \sum_{m=2}^N \sum_{j=2}^N \frac{(1-\lambda_m^n)(1-\lambda_j^n)}{(1-\lambda_m)(1-\lambda_j)} b_\alpha^{(m)} b_\gamma^{(j)} \right\} + p_{\alpha\beta} p_{\gamma\delta} \left\{ -n\pi_\alpha \pi_\gamma + \sum_{m=2}^N \frac{n-1-n\lambda_m + \lambda_m^n \pi_\alpha \pi_\gamma}{(1-\lambda_m)^2} (b_\alpha^{(m)} a_{\beta\gamma}^{(m)} + b_\gamma^{(m)} a_{\beta\alpha}^{(m)}) \right. \\ & \left. + \pi_\gamma (b_\alpha^{(n)} + a_{\beta\alpha}^{(n)}) \right] + \sum_{m=2}^N \frac{1-n\lambda_m^{n-1} + (n-1)\lambda_m^n}{(1-\lambda_m)^2} (b_\alpha^{(m)} a_{\beta\gamma}^{(m)} + b_\gamma^{(m)} a_{\beta\alpha}^{(m)}) \\ & + \sum_{m=2}^N \sum_{\substack{j \neq 1 \\ j \neq m}} \frac{b_\alpha^{(m)} a_{\beta\gamma}^{(j)} + b_\gamma^{(m)} a_{\beta\alpha}^{(j)}}{1-\lambda_j} \left[ \frac{1-\lambda_m^{n-1}}{1-\lambda_m} - \frac{\lambda_j(\lambda_m^{n-1} - \lambda_j^{n-1})}{\lambda_m - \lambda_j} \right]. \end{aligned} \quad (6.1.51)$$

Proof. Since

$$E[\tilde{x}_{\alpha\beta} \tilde{x}_{\gamma\delta}] = \sum_{u=1}^N p_u \sigma_{\alpha\beta\gamma\delta}^{(u,u)}, \quad (6.1.52)$$



equation (6.1.49) follows immediately from (6.1.27). By using (6.1.52) together with (6.1.50) and the spectral representation (6.1.29), equation (6.1.51) is obtained. Q.E.D.

**Corollary 6.1.9** Let  $\tilde{F}$  be an  $N \times N$  random matrix which has the Whittle-2 distribution with parameter  $(\underline{\pi}, n, P)$ , where  $P$  is ergodic and  $\underline{\pi}$  is the steady state distribution corresponding to  $P$ . Then

$$E[\tilde{F}_{ij}] = n \pi_{ij}^P, \quad i, j = 1, \dots, N \quad (6.1.53)$$

and, for  $a, \beta, \gamma, \delta = 1, \dots, N$ ,

$$\text{cov}[\tilde{F}_{ab}, \tilde{F}_{\gamma\delta}] =$$

$$n \pi_{ab}^P (\delta_{a\gamma} \delta_{b\delta} - n \pi_{\gamma\delta}^P) + \sum_{k=0}^{n-1} [(n-1-k) p_{ab} p_{\gamma\delta} (\pi_{ab}^{(k)} + \pi_{\gamma\delta}^{(k)})], \\ n=2, 3, \dots \quad (6.1.54a)$$

$$= \pi_{ab}^P (\delta_{a\gamma} \delta_{b\delta} - \pi_{\gamma\delta}^P), \quad n=1 \quad (6.1.54b)$$

If  $P$  has distinct eigenvalues, equation (6.1.54a) has the spectral representation,

$$\text{cov}[\tilde{F}_{ab}, \tilde{F}_{\gamma\delta}] =$$

$$n \pi_{ab}^P (\delta_{a\gamma} \delta_{b\delta} - \pi_{\gamma\delta}^P) + p_{ab} p_{\gamma\delta} \sum_{m=2}^N \frac{n-1-n\lambda_m + \lambda_m^n}{(1-\lambda_m)^2} [\pi_{ab}^{(m)} + \pi_{\gamma\delta}^{(m)}]. \\ n=2, 3, \dots \quad (6.1.55)$$

**Proof.** Since

$$\sum_{u=1}^N \pi_{uv}^{(k)} = \pi_1, \quad k=0, 1, 2, \dots \quad (6.1.56) \\ i=1, \dots, N$$

equation (6.1.53) follows immediately from (6.1.46). Using (6.1.56) in equation (6.1.49a),



$$\text{cov} [\tilde{\mathbb{P}}_{\alpha\beta}, \tilde{\mathbb{P}}_{\gamma\delta}] =$$

$$n\pi_{\alpha\beta}^p(\delta_{\alpha\gamma}\delta_{\beta\delta} - n\pi_{\gamma\delta}^p) + \sum_{k=1}^{n-1} [\pi_{\alpha\beta}^p \tilde{\mathbb{P}}_{\gamma\delta}(\beta, k) + \pi_{\gamma\delta}^p \tilde{\mathbb{P}}_{\alpha\beta}(\delta, k)], \\ n=2, 3, \dots \quad (6.1.57a)$$

$$= \pi_{\alpha\beta}^p (\delta_{\alpha\gamma}\delta_{\beta\delta} - \pi_{\gamma\delta}^p). \quad n=1 \quad (6.1.57b)$$

Noting that, if  $n > 1$ ,

$$\begin{aligned} \sum_{k=1}^{n-1} \tilde{\mathbb{P}}_{ij}(u, k) &= \sum_{k=1}^{n-1} \sum_{m=0}^{k-1} p_{ui}^{(m)} p_{ij}^{(m)} \\ &= \sum_{m=0}^{n-1} \sum_{k=m+1}^{n-1} p_{ui}^{(m)} p_{ij}^{(m)} \\ &= p_{ij} \sum_{m=0}^{n-1} (n-1-m) p_{ui}^{(m)}, \end{aligned} \quad (6.1.58)$$

equation (6.1.54) follows from (6.1.57).

If  $P$  has distinct eigenvalues, then, by (6.1.12a),

$$b_j^{(n)} = \sum_{u=1}^N \pi_{u\alpha}^{(n)} u_j^{(n)} = 0, \quad m=2, \dots, N \quad (6.1.59) \\ j=1, \dots, N$$

and, in this case, equation (6.1.51) reduces to (6.1.55). Q.E.D.

## 6.2 The Multivariate Beta Distribution.

In this section we consider the multivariate beta distribution, which is an extension of the beta distribution to  $N$  dimensions. There are several different generalizations of the beta distribution; this particular one is due to Mardia [30]. The moments of this distribution have been derived by Neumann [31], who also relates the multivariate beta distribution to the gamma distribution. Some of Neumann's results are presented here for the sake of completeness; the proofs, for the most part, are original.



6.2.1 The Multivariate Beta Density Function. The random stochastic vector,  $\tilde{p} = (\tilde{p}_1, \dots, \tilde{p}_N)$ , is said to have the multivariate beta distribution with parameter  $\underline{m}$  if  $\tilde{p}$  has the joint density function

$$f_{\tilde{p}}^{(N)}(\tilde{p} | \underline{m}) = B_N(\underline{m}) \prod_{i=1}^{N-1} p_i^{m_i-1} (1 - \sum_{k=1}^{N-1} p_k)^{m_N-1} \quad \text{for } \tilde{p} \in \mathcal{S}_{1,N}$$

$$= 0, \quad \text{elsewhere} \quad (6.2.1)$$

where  $\underline{m} = (m_1, \dots, m_N)$  with

$$m_i > 0, \quad i=1, \dots, N \quad (6.2.2)$$

and, if  $\Gamma(x)$  is the gamma function and

$$M = \sum_{i=1}^N m_i, \quad (6.2.3)$$

the normalizing constant is

$$B_N(\underline{m}) = \frac{\Gamma(M)}{\prod_{i=1}^N \Gamma(m_i)}. \quad (6.2.4)$$

It is to be noted that  $f_{\tilde{p}}^{(N)}(\tilde{p} | \underline{m})$  is the joint distribution of  $N-1$  of the elements of  $\tilde{p}$ , the  $N$ th element being determined by the constraint

$$\sum_{i=1}^N \tilde{p}_i = 1 \quad (6.2.5)$$

The following lemma provides an alternate representation of the normalizing constant,  $B_N(\underline{m})$ .

**Lemma 6.2.1** If  $B_N(\underline{m})$  is defined by (6.2.4), then

$$B_N(\underline{m}) = [B(m_1, \sum_{i=2}^N m_i) B(m_2, \sum_{i=3}^N m_i) \dots B(m_{N-1}, m_N)]^{-1}, \quad (6.2.6)$$

where  $B(n,n)$  is the beta function.



Proof.

$$\begin{aligned}
 B_N(n) &= \frac{\Gamma(n)}{\prod_{i=1}^N \Gamma(n_i)} \\
 &= \frac{\Gamma(n_1 + \sum_{i=2}^N n_i) \Gamma(n_2 + \sum_{i=3}^N n_i) \dots \Gamma(n_{N-1} + n_N)}{\Gamma(n_1) \Gamma(\sum_{i=2}^N n_i) \Gamma(n_2) \Gamma(\sum_{i=3}^N n_i) \dots \Gamma(n_{N-1}) \Gamma(n_N)} \\
 &= [B(n_1, \sum_{i=2}^N n_i) B(n_2, \sum_{i=3}^N n_i) \dots B(n_{N-1}, n_N)]^{-1} \quad (6.2.7)
 \end{aligned}$$

Q.E.D.

Since  $B(n, n) > 0$ , it is clear that

$$f_{\beta}^{(N)}(\underline{p} | \underline{n}) \geq 0, \quad \underline{p} \in \mathcal{S}_{1,N}$$

In the next theorem it is established that  $\int_{\mathcal{S}_{1,N}} f_{\beta}^{(N)}(\underline{p} | \underline{n}) d\underline{p} = 1$ . It then follows that the multivariate beta density function, as defined by equation (6.2.1), is a proper density function.

Theorem 6.2.2 If  $f_{\beta}^{(N)}(\underline{p} | \underline{n})$  is defined by equation (6.2.1), then

$$\int_{\mathcal{S}_{1,N}} f_{\beta}^{(N)}(\underline{p} | \underline{n}) d\underline{p} = 1, \quad (6.2.8)$$

where  $d\underline{p} = dp_1 \dots dp_{N-1}$ .

Proof. The theorem is proved by induction on  $N$ . For  $N=2$ ,  $f_{\beta}^{(2)}(\underline{p} | \underline{n})$  is the univariate beta density function and (6.2.8) holds. Assume (6.2.8) is true for  $N$ . Then,

$$\int_{\mathcal{S}_{1,N+1}} f_{\beta}^{(N+1)}(\underline{p} | \underline{n}) d\underline{p} = B_{N+1}(n) \int_{\mathcal{S}_{1,N+1}} \prod_{i=1}^N p_i^{n_i-1} (1 - \sum_{i=1}^N p_i)^{n_{N+1}-1} dp_1 \dots dp_N. \quad (6.2.9)$$



Let us make the integrand transformation

$$\underline{p}_1 = (p_1, \dots, p_{N-1}, 1 - \sum_{i=1}^{N-1} p_i) \quad (6.2.10a)$$

$$u = (u, 1-u), \quad (6.2.10b)$$

where

$$u = \frac{p_N}{1 - \sum_{i=1}^{N-1} p_i}.$$

The Jacobian is

$$\frac{dp_N}{du} = 1 - \sum_{i=1}^{N-1} p_i. \quad (6.2.11)$$

Noting that  $\underline{p}_1 \in \mathcal{S}_{1,N}$  and  $u \in \mathcal{S}_{1,2}$ , and letting

$$\underline{m}_1 = (m_1, \dots, m_{N-1}, m_N + m_{N+1})$$

$$\underline{m}_2 = (m_N, m_{N+1}),$$

equation (6.2.9) becomes

$$\begin{aligned} & \int_{\mathcal{S}_{1,N}} \prod_{i=1}^{N-1} p_i^{m_i-1} (1 - \sum_{i=1}^{N-1} p_i)^{m_N + m_{N+1}-1} dp_1 \dots dp_{N-1} \\ & \times \frac{1}{B(m_N, m_{N+1})} \int_0^1 u^{m_N-1} (1-u)^{m_{N+1}-1} du \\ & = \int_{\mathcal{S}_{1,N}} f_{\beta}^{(N)}(\underline{p}_1 | \underline{m}_1) dp_1 \int_{\mathcal{S}_{1,2}} f_{\beta}^{(2)}(\underline{u} | \underline{m}_2) d\underline{u} \\ & = 1. \end{aligned} \quad (6.2.12)$$

Q.E.D.



The method by which Theorem 6.2.2 was proved can be generalized to provide an identity which will be useful in subsequent proofs.

**Theorem 6.2.3** Let  $\tilde{\underline{p}}$  be a random stochastic vector with the multivariate beta distribution,  $f_{\beta}^{(N)}(\underline{p} | \underline{n})$ . Let  $\mu$  and  $v$  be positive integers such that  $\mu + v = N$  and make the transformation

$$\tilde{\underline{q}} = (\tilde{p}_1, \dots, \tilde{p}_{\mu}, 1 - \sum_{k=1}^{\mu} \tilde{p}_k) \quad (6.2.13a)$$

$$\tilde{\underline{u}} = (\tilde{u}_1, \dots, \tilde{u}_{v-1}, 1 - \sum_{i=1}^{v-1} \tilde{u}_i) \quad (6.2.13b)$$

where

$$\tilde{u}_i = \frac{\tilde{p}_{\mu+i}}{1 - \sum_{k=1}^{\mu} \tilde{p}_k} . \quad i=1, \dots, v-1$$

Let

$$\underline{u}_1 = (u_1, \dots, u_{\mu}, \sum_{k=1}^{\mu} \frac{u}{\mu+k}) \quad (6.2.14a)$$

$$\underline{u}_2 = (u_{\mu+1}, \dots, u_N). \quad (6.2.14b)$$

Then the joint distribution of  $(\tilde{\underline{q}}, \tilde{\underline{u}})$  is

$$D(\underline{q}, \underline{u} | \underline{n}_1, \underline{n}_2) = f_{\beta}^{(\mu+1)}(\underline{q} | \underline{n}_1) f_{\beta}^{(v)}(\underline{u} | \underline{n}_2). \quad (6.2.15)$$

**Proof.** Since  $0 \leq \tilde{u}_i \leq 1$  ( $i=1, \dots, v-1$ ), the range sets of  $\tilde{\underline{q}}$  and  $\tilde{\underline{u}}$  are, respectively,  $\mathcal{S}_{1, \mu+1}$  and  $\mathcal{S}_{1, v}$ . The Jacobian of the transformation (6.2.13) is

$$J \left( \frac{p_{\mu+1}, \dots, p_{N-1}}{u_1, \dots, u_{v-1}} \right) = \left| \frac{\partial p_{\mu+1}}{\partial u_j} \right| = \left( 1 - \sum_{k=1}^{\mu} p_k \right)^{v-1}. \quad (6.2.16)$$

Making the transformation (6.2.13) in (6.2.1),



$$D(\underline{q}, \underline{u} | \underline{s}_1, \underline{s}_2) = B_{\mu+1}(s_1) B_{\nu}(s_2) \prod_{i=1}^{\mu} p_i^{u_i-1} \left(1 - \sum_{k=1}^N p_k\right)^{\nu-1}$$

$$\times \prod_{j=1}^{\nu-1} u_j^{m_{\mu+j}-1} \left(1 - \sum_{k=1}^{\nu-1} u_k\right)^{m_N-1}$$

$$= f_{\beta}^{(\mu+1)}(\underline{q} | \underline{s}_1) f_{\beta}^{(\nu)}(\underline{u} | \underline{s}_2). \quad (6.2.17)$$

Q.E.D.

### 6.2.2 The Multivariate Beta Distribution Function.

If the random stochastic vector  $\underline{p}$  has the multivariate beta distribution with parameter  $\underline{n}$ , then the multivariate beta distribution function is denoted  $F_{\beta}^{(N)}(\underline{p} | \underline{n})$ , where, if  $\underline{p} = (p_1, \dots, p_N)$ ,

$$F_{\beta}^{(N)}(\underline{p} | \underline{n}) = P [\underline{p}_1 \leq p_1, \dots, \underline{p}_{N-1} \leq p_{N-1}]$$

$$= \int_0^{p_1} \dots \int_0^{p_{N-1}} f_{\beta}^{(N)}(\underline{q} | \underline{n}) d\underline{q}. \quad (6.2.18)$$

As is the case with the beta distribution function, there is no closed expression for equation (6.2.18). We can, however, express  $F_{\beta}^{(N)}(\underline{p} | \underline{n})$  as an  $(N-1)$ -fold infinite sum. The case  $N = 3$  is illustrated here.

For notational simplicity, let the parameter vector be

$$\underline{n} = (\alpha, \beta, \gamma). \quad (6.2.19)$$

Then

$$F_{\beta}^{(3)}(\underline{p} | \underline{n}) = B_3(n) \int_0^{p_1} \int_0^{p_2} q_1^{\alpha-1} q_2^{\beta-1} (1-q_1-q_2)^{\gamma-1} dq_1 dq_2. \quad (6.2.20)$$

Prob 1,3



If  $0 < p_1 + p_2 < 1$ , we have  $0 < q_1 + q_2 < 1$  in the range of integration and

$$(1 - q_1 - q_2)^{\gamma-1} = \sum_{k=0}^{\infty} \binom{\gamma-1}{k} (-1)^k (q_1 + q_2)^k, \quad (6.2.21)$$

the series converging uniformly. Since  $k$  is a positive integer, we may expand  $(q_1 + q_2)^k$  in a finite binomial series,

$$(q_1 + q_2)^k = \sum_{v=0}^k \binom{k}{v} q_1^{k-v} q_2^v. \quad (6.2.22)$$

Thus,

$$\begin{aligned} r_p^{(3)}(\underline{p} | \underline{q}) &= B_3(\underline{\alpha}) \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{\gamma-1}{k} \binom{k}{v} (-1)^k \int_0^{p_1} q_1^{k+\alpha-v-1} dq_1 \int_0^{p_2} q_2^{\beta+v-1} dq_2 \\ &= B_3(\underline{\alpha}) \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{\gamma-1}{k} \binom{k}{v} (-1)^k \frac{p_1^{k+\alpha-v}}{k+\alpha-v} \frac{p_2^{\beta+v}}{\beta+v}. \end{aligned} \quad (6.2.23)$$

Let  $(x)_k$  denote the hypergeometric coefficient

$$\begin{aligned} (x)_k &= x(x+1) \dots (x+k-1) & k=1, 2, \dots \\ &= 1, & k=0 \end{aligned} \quad (6.2.24)$$

where  $x$  is any real number. Then

$$(-1)^k \binom{\gamma-1}{k} \binom{k}{v} = \frac{(\gamma)_k}{(k-v)! v!} \quad (6.2.25)$$

and (6.2.23) becomes

$$B_3(\underline{\alpha}) \frac{\alpha}{p_1} \frac{\beta}{p_2} \sum_{k=0}^{\infty} \sum_{v=0}^k \frac{(\gamma)_k}{(k+\alpha-v)(\beta+v)} \frac{p_1^{k+\alpha-v}}{(k-v)!} \frac{p_2^v}{v!} \quad (6.2.26)$$

Reversing the order of summation and noting that

$$\alpha + k = \frac{\alpha(\alpha + 1)}{k}$$



$$\beta + v = \frac{\beta(\beta + 1)}{(\beta)_v} v$$

we have

$$F_{\beta}^{(3)}(p | \underline{z}) = B_3(\underline{z}) \frac{p^{\alpha} p^{\beta}}{\alpha! \beta!} \sum_{v=0}^{\infty} \sum_{k=0}^{\infty} \frac{(1-v)_{k+v} (\alpha)_k (\beta)_v}{(v+1)_k (\beta+1)_v} \frac{p_1^k}{k!} \frac{p_2^k}{v!} \quad (6.2.27)$$

The double infinite series of (6.2.27) is  $F_2(1-\gamma, \alpha, \beta, \alpha+1, \beta+1; p_1, p_2)$ ,

Appell's second hypergeometric function of two variables [2]. Appell has shown that the double series of (6.2.27) converges absolutely whenever  $p_1 + p_2 < 1$ . Thus, we have

$$F_{\beta}^{(3)}(p | \underline{z}) = B_3(\underline{z}) \frac{p_1^{\alpha} p_2^{\beta}}{\alpha! \beta!} F_2(1-\gamma, \alpha, \beta, \alpha+1, \beta+1; p_1, p_2). \quad \begin{matrix} p \in \mathcal{P}_{1,3} \\ p_1 + p_2 < 1 \end{matrix} \quad (6.2.28)$$

The question of convergence of (6.2.27) on the boundary of the region  $p_1 + p_2 \leq 1$  has not yet been resolved. We note, however, that equation (6.2.23) remains valid when  $p_1 + p_2 = 1$ .

### 6.2.3 The Nonstandardized Multivariate Beta Distribution. Let

$\tilde{\underline{q}} = (\tilde{q}_1, \dots, \tilde{q}_N)$  be a random vector with range set

$$R_N(a) = \left\{ \underline{q} \mid 0 \leq q_i \leq a \ (i=1, \dots, N), \sum_{i=1}^N q_i = a \right\}, \quad (6.2.29)$$

where  $a > 0$ . The vector  $\tilde{\underline{q}}$  is said to have the nonstandardized multivariate beta distribution with parameter  $(\alpha, \beta)$  if  $\tilde{\underline{q}}$  has the density function

$$f_{\beta^N}^{(N)}(\underline{q} | \alpha, \underline{z}) = \frac{B_N(\underline{z})}{a^{N-1}} \prod_{i=1}^{N-1} q_i^{\alpha_i - 1} (a - \sum_{k=i+1}^{N-1} q_k)^{\alpha_{N-i}} \quad (6.2.30)$$

$\begin{matrix} q \in R_N(a) \\ = 0, \end{matrix}$  elsewhere



where  $\underline{m} = (m_1, \dots, m_N)$  and

$$a > 0 \quad (6.2.31a)$$

$$m_i > 0 \quad i=1, \dots, N \quad (6.2.31b)$$

$$M = \sum_{i=1}^N m_i. \quad (6.2.31c)$$

The nonstandardized multivariate beta distribution is obtained from equation (6.2.1) by making the transformation

$$\underline{\tilde{q}} = a\underline{q}, \quad (6.2.32)$$

which has the Jacobian

$$J \left( \frac{p_1, \dots, p_{N-1}}{q_1, \dots, q_{N-1}} \right) = \frac{1}{a^{N-1}}. \quad (6.2.33)$$

It then follows that  $f_{\beta^*}^{(N)}(\underline{q} | a, \underline{m}) \geq 0$  and

$$\int_{R_{N-1}(4)} f_{\beta^*}^{(N)}(\underline{q} | a, \underline{m}) dq_1 \dots dq_{N-1} = 1.$$

**6.2.4 Marginal and Conditional Distributions.** Suppose that  $\tilde{\underline{p}}$  is an  $N$ -dimensional random stochastic vector with the density function  $f_{\beta}^{(N)}(\underline{p} | \beta)$ . We now show that the marginal and conditional distributions of  $v$  of the elements of  $\tilde{\underline{p}}$  ( $v=1, 2, \dots, N-2$ ) are, respectively, multivariate beta and nonstandardised multivariate beta. It is assumed, without loss of generality, that the elements of interest are  $\tilde{p}_1, \dots, \tilde{p}_v$ .

Let

$$\tilde{\underline{q}}(v) = (\tilde{p}_1, \dots, \tilde{p}_v, 1 - \sum_{i=1}^v \tilde{p}_i) \quad (6.2.34)$$

and

$$\tilde{\underline{s}}(v) = (\tilde{p}_1, \dots, \tilde{p}_v, 1 - \tilde{s}(v) - \sum_{i=1}^v \tilde{p}_i) \quad (6.2.35)$$



where

$$\tilde{b}(v) = \sum_{i=v+1}^{N-1} \tilde{p}_i. \quad (6.2.36)$$

Also, let

$$\tilde{\mathbf{t}}(v) = (\tilde{p}_{v+1}, \dots, \tilde{p}_{N-1}). \quad (6.2.37)$$

**Theorem 6.2.4** If the random  $N$ -dimensional stochastic vector  $\tilde{\mathbf{p}}$  has the multivariate beta distribution with parameter  $\underline{\alpha}$ , then, for  $v=1, \dots, N-2$ , the marginal distribution of  $(\tilde{p}_1, \dots, \tilde{p}_v)$  is multivariate beta,

$$P(p_1, \dots, p_v | \underline{\alpha}) = f_{\beta^v}^{(v+1)}(\mathbf{q}(v) | \underline{\alpha}(v)), \quad (6.2.38)$$

and the conditional distribution of  $(\tilde{p}_1, \dots, \tilde{p}_v)$ , given that

$(\tilde{p}_{v+1}, \dots, \tilde{p}_{N-1}) = (p_{v+1}, \dots, p_{N-1})$ , is nonstandardised multivariate beta,

$$D(p_1, \dots, p_v | \underline{\alpha}, \tilde{\mathbf{t}}^*(v)) = f_{\beta^v}^{(v+1)}(\mathbf{t}(v) | 1-b(v), \underline{\alpha}^*(v)), \quad (6.2.39)$$

where

$$\underline{\alpha}(v) = (n_1, \dots, n_v, \sum_{i=v+1}^N n_i) \quad (6.2.40)$$

and

$$\underline{\alpha}^*(v) = (n_1, \dots, n_v, n_N). \quad (6.2.41)$$

**Proof.** Let

$$\tilde{\mathbf{u}} = (\tilde{u}_1, \dots, \tilde{u}_{N-v-1}, 1 - \sum_{i=1}^{N-v-1} \tilde{u}_i)$$

where

$$\tilde{u}_i = \frac{\tilde{p}_{v+1}}{1 - \sum_{k=1}^v \tilde{p}_k} \quad i=1, \dots, N-v-1$$

and let

$$\mathbf{R}_2 = (n_{v+1}, \dots, n_N).$$



Then, by Theorem 6.2.3, the joint density function of  $(\tilde{q}(v), \tilde{u})$  is

$$D(q(v), u | \underline{\mu}(v), \underline{B}_2) = f_{\beta}^{(v+1)}(q(v) | \underline{\mu}(v)) f_{\beta}^{(N-v)}(u | \underline{B}_2). \quad (6.2.42)$$

The marginal density function of  $\tilde{q}(v)$  is, therefore,

$$\int_{\underline{L}_{1,N-v}} D(q(v), u | \underline{\mu}(v), \underline{B}_2) du_1 \dots du_{N-v-1} = f_{\beta}^{(v+1)}(q(v) | \underline{\mu}(v)). \quad (6.2.43)$$

If  $(\tilde{p}_{v+1}, \dots, \tilde{p}_{N-1}) = (p_{v+1}, \dots, p_{N-1})$  then the random variables  $\tilde{p}_1, \dots, \tilde{p}_v$  are constrained by

$$0 \leq \sum_{i=1}^v p_i \leq 1-b(v) \quad \text{ini, ..., v} \quad (6.2.44)$$

and the conditional density function of  $(\tilde{p}_1, \dots, \tilde{p}_v)$  has the kernel

$$\prod_{i=1}^v p_i^{m_i-1} (1-b(v) - \sum_{i=1}^v p_i)^{n_{N-i}-1} \quad (6.2.45)$$

which is the kernel of the nonstandardized multivariate beta density function.

$$f_{\beta}^{(v+1)}(q(v) | 1-b(v), \underline{\mu}^*(v)). \quad \text{Q.E.D.}$$

**6.2.5 Moment Formulas.** The moments of the multivariate beta distribution are most easily derived as a special case of the moments of the matrix beta distribution, to be considered in Section 6.3. These results are stated here and proved in the next section.

Let the random stochastic vector  $\tilde{p} = (\tilde{p}_1, \dots, \tilde{p}_N)$  have the multivariate beta distribution with parameter  $\underline{n} = (n_1, \dots, n_N)$ . Then, if

$$M = \sum_{i=1}^N n_i,$$

$$E[\tilde{p}_i] = \tilde{p}_i(M) = \frac{n_i}{M}, \quad i=1, \dots, N \quad (6.2.46)$$

$$\text{var}(\tilde{p}_i) = \frac{n_i(M-n_i)}{M^2(M+1)}, \quad i=1, \dots, N \quad (6.2.47)$$



and

$$\text{cov}[\tilde{p}_i, \tilde{p}_j] = \frac{-\frac{m_i m_j}{N^2(N+1)}}{\frac{m_i m_j}{N^2(N+1)}}, \quad i \neq j, \dots, N \quad (6.2.48)$$

All higher moments can be computed from the recurrence relation

$$E\left[\prod_{i=1}^N \tilde{p}_i^{v_i} | \underline{m}\right] = \tilde{p}_a(\underline{m}) E\left[\tilde{p}_a^{v_a-1} \prod_{i \neq a} \tilde{p}_i^{v_i} | T_a(\underline{m})\right], \quad (6.2.49)$$

where the  $v_i$  are nonnegative integers,  $a$  is any index such that  $v_a > 0$ , and  $T_a(\underline{m})$  is the vector  $\underline{m}$  with the element  $m_a$  increased by unity.

In matrix form, (6.2.46) - (6.2.48) can be summarized as the mean vector,

$$E[\tilde{\underline{p}}] = \frac{1}{N} \underline{m} \quad (6.2.50)$$

and the variance-covariance matrix,

$$V[\tilde{\underline{p}}] = \frac{1}{N^2(N+1)} [\underline{M} - \underline{m} \underline{m}^T], \quad (6.2.51)$$

where  $\underline{M}$  is the  $N \times N$  diagonal matrix  $[m_{ij}]$ .

Using (6.2.32), it is easily seen that, if  $\tilde{\underline{q}}$  has the nonstandardized multivariate beta density function  $f_{\beta^*}^{(N)}(\underline{q} | a, b)$ , then

$$E[\tilde{\underline{q}}] = \frac{a}{N} \underline{m} \quad (6.2.52)$$

and

$$V[\tilde{\underline{q}}] = \frac{a^2}{N^2(N+1)} [\underline{M} - \underline{m} \underline{m}^T]. \quad (6.2.53)$$

### 6.3 The Matrix Beta Distribution.

The  $K \times N$  random generalized stochastic matrix,  $\tilde{\underline{\underline{P}}} = [\tilde{p}_{ij}^k]$ , is said to have the matrix beta distribution with parameter  $\underline{\underline{M}} = [m_{ij}^k]$  if  $\tilde{\underline{\underline{P}}}$  has the joint density function

$$f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{M}}) = \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^{K_i} B_N(m_{i-1}^k) (p_{ij}^k)^{m_{ij}^k - 1} \quad \underline{\underline{P}} \in \mathcal{S}_{K,N} \\ = 0, \quad \text{elsewhere} \quad (6.3.1)$$



where  $\underline{\underline{M}}$  is a  $K \times N$  matrix such that

$$m_{ij}^k > 0, \quad k=1, \dots, K_1 \quad (6.3.2)$$

$$1, j=1, \dots, N$$

and  $E_N(\underline{\underline{m}}_1^k)$  is defined by (6.2.4). The generic row of  $\underline{\underline{M}}$  is denoted  $\underline{\underline{m}}_1^k$ , an  $N$ -dimensional vector. The total number of rows of both  $\underline{\underline{P}}$  and  $\underline{\underline{M}}$  is  $K = \sum_{i=1}^N K_i$ . To be quite general, we admit the possibility that  $K_i = 0$  for some  $i$ .

The matrix beta distribution is the joint distribution of  $K(N-1)$  random variables,  $\hat{p}_{ij}^k$ . The remaining  $K$  elements of  $\underline{\underline{P}}$  are determined by the relations

$$\sum_{j=1}^N \hat{p}_{ij}^k = 1 \quad k=1, \dots, K_1 \quad (6.3.3)$$

$$i=1, \dots, N$$

Inspection of equation (6.3.1) shows that the matrix beta density function is the product of  $K$  multivariate beta density functions,

$$f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{M}}) = \prod_{i=1}^N \prod_{k=1}^{K_i} f_B^{(N)}(p_{ij}^k | m_{ij}^k). \quad (6.3.4)$$

It follows that  $f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{M}}) \geq 0$  and

$$\int_{\mathcal{S}_{K,N}} f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{M}}) d\underline{\underline{P}} = 1, \quad (6.3.5)$$

$$\text{where } d\underline{\underline{P}} = \prod_{i=1}^N \prod_{k=1}^{K_i} \prod_{j=1}^{N-1} dp_{ij}^k.$$

The family of matrix beta distributions is a family indexed by the parameter  $\underline{\underline{M}} \in \Psi_{K,N}$ , where the admissible parameter set is

$$\Psi_{K,N} = \left\{ \underline{\underline{M}} \mid \underline{\underline{M}} \text{ is } K \times N, \quad m_{ij}^k > 0 \quad (k=1, \dots, K_1; i, j=1, \dots, N) \right\}, \quad (6.3.6)$$

the positive orthant of  $E_{KN}$ . Since  $f_{MB}^{(K,N)}(\underline{\underline{P}} | \underline{\underline{M}})$  is a continuous function of  $\underline{\underline{M}}$  for all  $\underline{\underline{P}} \in \mathcal{S}_{K,N}$ , Corollary 2.4.5 implies that the family of matrix



beta distributions is continuous in  $\underline{M}$ . In Theorem 2.2.1 it was shown that the matrix beta distribution is the natural conjugate distribution for a Markov chain which is observed under the consecutive sampling rule. It then follows that the family of matrix beta distributions is closed under consecutive sampling. This property is used in the following theorem to derive the moments of the matrix beta distribution.

**Theorem 5.3.1** Let  $\underline{\underline{P}} = [p_{ij}^k]$  be a random generalized stochastic matrix which has the matrix beta distribution with parameter  $\underline{M} = [m_{ij}^k]$ .

Then,

$$E[p_{ij}^k] = p_{ij}^k(\underline{M}) = \frac{m_{ij}^k}{M_i^k} \quad \begin{matrix} k=1, \dots, K \\ i, j=1, \dots, N \end{matrix} \quad (6.3.7)$$

$$\text{var}[p_{ij}^k] = \frac{m_{ij}^k(M_i^k - m_{ij}^k)}{(M_i^k)^2(M_i^k + 1)}, \quad \begin{matrix} k=1, \dots, K \\ i, j=1, \dots, N \end{matrix} \quad (6.3.8)$$

$$\text{cov}[p_{\alpha\beta}^k, p_{\gamma\delta}^j] = \frac{-m_{\alpha\beta}^k m_{\gamma\delta}^j}{(M_\alpha^k)^2(M_\alpha^k + 1)}, \quad \begin{matrix} j=i=1, \dots, K \\ \alpha=\gamma=1, \dots, N \\ \beta, \delta=1, \dots, N \\ \beta \neq \delta \end{matrix} \quad (6.3.9a)$$

$$= 0, \quad j \neq k \text{ or } \alpha \neq \gamma \quad (6.3.9b)$$

where

$$M_i^k = \sum_{j=1}^N m_{ij}^k. \quad \begin{matrix} k=1, \dots, K \\ i=1, \dots, N \end{matrix} \quad (6.3.10)$$

**Proof.** Let  $T_{ij}^k(\underline{M})$  be the matrix  $\underline{M}$  with the element  $m_{ij}^k$  increased by unity. Then

$$\begin{aligned} E[p_{ij}^k] &= \int_{\Omega_{K,N}} p_{ij}^k f_{MP}^{(K,N)}(\underline{\underline{P}} | \underline{M}) d\underline{\underline{P}} \\ &= \frac{\Gamma(M_i^k)}{\prod_{a=1}^N \Gamma(m_{ia}^k)} \cdot \frac{\prod_{a=1}^N \Gamma(m_{ia}^k + \delta_{aj})}{\Gamma(M_i^k + 1)} \int_{\Omega_{K,N}} f_{MP}^{(K,N)}(\underline{\underline{P}} | T_{ij}^k(\underline{M})) d\underline{\underline{P}} \\ &= \frac{m_{ij}^k}{M_i^k} \end{aligned} \quad (6.3.11)$$



For  $j \neq k$  or  $\alpha \neq \gamma$ ,  $\tilde{p}_{\alpha\beta}^k$  and  $\tilde{p}_{\gamma\delta}^j$  are independent random variables and  $\text{cov}[\tilde{p}_{\alpha\beta}^k, \tilde{p}_{\gamma\delta}^j] = 0$ . If  $j = k$  and  $\alpha = \gamma$ , Lemma 2.3.2 yields

$$\begin{aligned} E[\tilde{p}_{\alpha\beta}^k \tilde{p}_{\alpha\delta}^k] &= \int_{\mathcal{S}_{K,N}} p_{\alpha\beta}^k p_{\alpha\delta}^k f_{M\beta}^{(K,N)}(\underline{\xi}|\underline{m}) d\underline{\xi} \\ &= \tilde{p}_{\alpha\beta}^k(\underline{m}) \tilde{p}_{\alpha\delta}^k(T_{\alpha\beta}^k(\underline{m})) \\ &= \frac{\frac{m_{\alpha\beta}^k}{m_{\alpha\beta}^k + 1}}{\frac{m_{\alpha}^k}{m_{\alpha}^k + 1}}, \quad \beta \neq \delta \end{aligned} \quad (6.3.12a)$$

$$= \frac{\frac{m_{\alpha\beta}^k}{m_{\alpha\beta}^k + 1}}{\frac{m_{\alpha}^k}{m_{\alpha}^k + 1}}, \quad \beta = \delta \quad (6.3.12b)$$

from which equations (6.3.8) and (6.3.9a) follow. Q.E.D.

By writing equation (6.3.8) as

$$\text{var}[\tilde{p}_{ij}^k] = \frac{\tilde{p}_{ij}^k (1 - \tilde{p}_{ij}^k)}{m_i^k + 1}, \quad (6.3.13)$$

where  $\tilde{p}_{ij}^k = \tilde{p}_{ij}^k(\underline{m})$ , it is seen that

$$0 \leq \text{var}[\tilde{p}_{ij}^k] < \frac{1}{4}. \quad \begin{matrix} i=1, \dots, K_N \\ j=1, \dots, K_N \end{matrix} \quad (6.3.14)$$

Similarly, equation (6.3.9a) can be written as

$$\text{cov}[\tilde{p}_{\alpha\beta}^k, \tilde{p}_{\alpha\delta}^k] = -\frac{\tilde{p}_{\alpha\beta}^k \tilde{p}_{\alpha\delta}^k}{m_{\alpha}^k + 1} \quad (6.3.15)$$

and, since  $\tilde{p}_{\alpha\beta}^k \leq 1 = \tilde{p}_{\alpha\beta}^k$ , we have the bounds

$$-\text{var}[\tilde{p}_{\alpha\beta}^k] \leq \text{cov}[\tilde{p}_{\alpha\beta}^k, \tilde{p}_{\alpha\delta}^k] \leq 0 \quad (6.3.16a)$$

$$-\text{var}[\tilde{p}_{\alpha\delta}^k] \leq \text{cov}[\tilde{p}_{\alpha\beta}^k, \tilde{p}_{\alpha\delta}^k] \leq 0 \quad (6.3.16b)$$

for  $k = 1, \dots, K_\alpha$ ;  $\alpha, \beta, \delta = 1, \dots, N$ .



When  $\tilde{\underline{\Theta}}$  has the matrix beta distribution the rows of  $\tilde{\underline{\Theta}}$  are mutually independent; thus, the general joint moment is

$$E \left[ \prod_{i=1}^N \prod_{j=1}^N \prod_{k=1}^K (\tilde{p}_{ij}^k)^{v_{ij}^k} \right] = \prod_{i=1}^N \prod_{k=1}^K E \left[ \prod_{j=1}^N (\tilde{p}_{ij}^k)^{v_{ij}^k} \right], \quad (6.3.17)$$

where the  $v_{ij}^k$  are non-negative integers. Let

$$E \left[ \prod_{j=1}^N (\tilde{p}_{ij}^k)^{v_{ij}^k} \mid \underline{\mathcal{M}} \right] = \int_{\Delta_{K,N}} \prod_{j=1}^N (\tilde{p}_{ij}^k)^{v_{ij}^k} \pi_{\underline{\mathcal{M}}}^{(K,N)}(\underline{\varphi} \mid \underline{\mathcal{M}}) d\underline{\varphi}. \quad (6.3.18)$$

The following theorem provides a recursive formula for computing this expectation.

**Theorem 6.3.2** If the  $K \times N$  random matrix  $\tilde{\underline{\Theta}}$  has the matrix beta distribution with parameter  $\underline{\mathcal{M}}$ , then

$$E \left[ \prod_{j=1}^N (\tilde{p}_{ij}^k)^{v_{ij}^k} \mid \underline{\mathcal{M}} \right] = \tilde{p}_{ia}^k(\underline{\mathcal{M}}) E \left[ (\tilde{p}_{ia}^k)^{v_{ia}^k - 1} \prod_{j \neq a} (\tilde{p}_{ij}^k)^{v_{ij}^k} \mid \tilde{\underline{\Theta}}_{ia}^k(\underline{\mathcal{M}}) \right],$$

$\begin{matrix} k=1, \dots, K \\ i=1, \dots, N \end{matrix}$  (6.3.19)

where the  $v_{ij}^k$  are nonnegative integers and  $a$  is any index such that  $\tilde{p}_{ia}^k > 0$ .

**Proof.** The theorem follows immediately by applying Lemma 2.3.2 to equation (6.3.18). Q.E.D.

Since the multivariate beta distribution is a special case of the matrix beta distribution in which  $K = 1$ , we immediately have the following corollary.

**Corollary 6.3.3** Let the random stochastic vector  $\tilde{\underline{p}} = (\tilde{p}_1, \dots, \tilde{p}_N)$  have the multivariate beta distribution with parameter  $\underline{\alpha} = (\alpha_1, \dots, \alpha_N)$ .

Then, if  $M = \sum_{i=1}^N \alpha_i$ ,



$$E[\tilde{p}_1] = \tilde{p}_1(\mathbf{n}) = \frac{n_1}{N}, \quad i=1, \dots, N \quad (6.3.20)$$

$$\text{var}[\tilde{p}_1] = \frac{n_1(N-n_1)}{N^2(N+1)}, \quad i=1, \dots, N \quad (6.3.21)$$

$$\text{cov}[\tilde{p}_1, \tilde{p}_j] = \frac{-n_1 n_j}{N^2(N+1)}, \quad i \neq j \quad (6.3.22)$$

and

$$E\left[\prod_{i=1}^N \tilde{p}_i^{v_i} | \mathbf{n}\right] = \tilde{p}_a(\mathbf{n}) E\left[\tilde{p}_a^{v_a-1} \prod_{i \neq a} \tilde{p}_i^{v_i} | T_a(\mathbf{n})\right] \quad (6.3.23)$$

where the  $v_i$  are nonnegative integers,  $a$  is any index such that  $v_a > 0$ , and  $T_a(\mathbf{n})$  is the vector  $\mathbf{n}$  with the element  $n_a$  increased by unity.

Before considering the marginal and conditional distributions of submatrices of  $\tilde{\underline{P}}$ , it is necessary to define the nonstandardized matrix beta distribution. Let

$$\underline{a} = (a_1^1, \dots, a_1^{K_1}, a_2^1, \dots, a_N^{K_N}) \quad (6.3.24)$$

be a  $K$ -dimensional vector, where

$$a_1^k > 0, \quad k=1, \dots, K_1 \quad (6.3.25)$$

Let  $\tilde{\underline{P}} = [\tilde{p}_{ij}^k]$  be a  $K \times N$  random matrix with range set

$$R_{K,N}(\underline{a}) =$$

$$\left\{ \underline{P} \mid \underline{P} \text{ is } K \times N, p_{ij}^k \geq 0, \sum_{j=1}^N p_{ij}^k = a_1^k, (i=1, \dots, K_1; 1, j=1, \dots, N) \right\}. \quad (6.3.26)$$

Let  $\underline{M} = [m_{ij}^k]$  be a  $K \times N$  matrix of positive elements. Then  $\tilde{\underline{P}}$  is said to have the nonstandardized matrix beta distribution with parameter  $(\underline{a}, \underline{M})$  if  $\tilde{\underline{P}}$  has the joint density function



$$f_{\underline{M}\beta}^{(K, N)}(\underline{\Phi} \mid \underline{a}, \underline{m}) = \prod_{i=1}^N \prod_{k=1}^{K_i} f_{\beta^k}^{(N)}(p_i^k \mid a_i^k, m_i^k), \quad (6.3.27)$$

where  $p_i^k$  and  $m_i^k$  are generic rows of  $\underline{\Phi}$  and  $\underline{M}$ , respectively, and

$f_{\beta^k}^{(N)}(p_i^k \mid a_i^k, m_i^k)$  is the nonstandardized multivariate beta distribution defined by equation (6.2.30).

We now consider the marginal and conditional distributions of any  $\rho \times v$  submatrix of  $\widetilde{\underline{\Phi}}$  when  $\widetilde{\underline{\Phi}}$  has the matrix beta distribution. To simplify the notation, assume that the elements of  $\widetilde{\underline{\Phi}}$  and  $\underline{M}$  have been relabelled so that  $\widetilde{\underline{\Phi}} = [\widetilde{p}_{ij}]$  and  $\underline{M} = [m_{ij}]$  ( $i=1, \dots, K$ ;  $j=1, \dots, N$ ), and that the submatrix of  $\widetilde{\underline{\Phi}}$  which is of interest consists of the elements  $\widetilde{p}_{ij}$  ( $i=1, \dots, \rho$ ;  $j=1, \dots, v$ ), where  $\rho \in \{1, \dots, K\}$  and  $v \in \{1, \dots, N-1\}$ . Define the  $\rho \times (v+1)$  matrix

$$\widetilde{\underline{\Phi}}_{\rho v} = \begin{bmatrix} \widetilde{p}_{11} & \dots & \widetilde{p}_{1v} & 1 - \sum_{j=1}^v \widetilde{p}_{1j} \\ \dots & & & \\ \widetilde{p}_{\rho 1} & \dots & \widetilde{p}_{\rho v} & 1 - \sum_{j=1}^v \widetilde{p}_{\rho j} \end{bmatrix} \quad (6.3.28)$$

and the  $\rho \times (v+1)$  matrix

$$\widetilde{\underline{\Phi}}_{\rho v}^* = \begin{bmatrix} \widetilde{p}_{11} & \dots & \widetilde{p}_{1v} & 1 - \widetilde{b}_1(v) - \sum_{j=1}^v \widetilde{p}_{1j} \\ \dots & & & \\ \widetilde{p}_{\rho 1} & \dots & \widetilde{p}_{\rho v} & 1 - \widetilde{b}_\rho(v) - \sum_{j=1}^v \widetilde{p}_{\rho j} \end{bmatrix} \quad (6.3.29)$$

where

$$\widetilde{b}_i(v) = \sum_{j=v+1}^{N-1} \widetilde{p}_{ij}. \quad i=1, \dots, \rho \quad (6.3.30)$$

Correspondingly, we define the  $\rho \times (v+1)$  parameter matrices,



$$H_{\rho v} = \begin{bmatrix} m_{11} & \cdots & m_{1v} & \sum_{j=v+1}^N m_{1j} \\ \vdots & \ddots & & \\ m_{\rho 1} & \cdots & m_{\rho v} & \sum_{j=v+1}^N m_{\rho j} \end{bmatrix} \quad (6.3.31)$$

$$\tilde{H}_{\rho v}^* = \begin{bmatrix} m_{11} & \cdots & m_{1v} & m_{1N} \\ \vdots & \ddots & & \\ m_{\rho 1} & \cdots & m_{\rho v} & m_{\rho N} \end{bmatrix}. \quad (6.3.32)$$

**Theorem 6.3.4.** Let the  $K \times N$  random generalised stochastic matrix  $\tilde{\mathbf{P}}$  have the matrix beta distribution with parameter  $\underline{m}$ . Then, for  $\rho = 1, \dots, K$  and  $v = 1, \dots, N-1$ , the marginal joint distribution of  $(\tilde{p}_{11}, \dots, \tilde{p}_{1v}, \tilde{p}_{21}, \dots, \tilde{p}_{\rho v})$  is matrix beta,

$$D(p_{11}, \dots, p_{\rho v} | \underline{m}) = f_{M^*}^{(\rho, v+1)}(\tilde{p}_{\rho v} | H_{\rho v}), \quad (6.3.33)$$

and the conditional joint distribution of  $(\tilde{p}_{11}, \dots, \tilde{p}_{\rho v})$ , given that  $(\tilde{p}_{1,v+1}, \dots, \tilde{p}_{1,N-1}, \tilde{p}_{2,v+1}, \dots, \tilde{p}_{K,N-1}) = (p_{1,v+1}, \dots, p_{1,N-1}, p_{2,v+1}, \dots, p_{K,N-1})$ , is nonstandardized matrix beta,

$$D(p_{11}, \dots, p_{\rho v} | \underline{m}, p_{1,v+1}, \dots, p_{K,N-1}) = f_{M^*}^{(\rho, v+1)}(\tilde{p}_{\rho v} | h^*(v), \tilde{H}_{\rho v}^*) \quad (6.3.34)$$

where

$$h^*(v) = (1 - b_1(v), \dots, 1 - b_\rho(v)). \quad (6.3.35)$$

**Proof.** The theorem follows immediately from Theorem 6.2.4, upon noting that the matrix beta density function is the product of  $K$  multivariate beta density functions.



#### 6.4 Extended Natural Conjugate Distributions.

If  $\tilde{\underline{\Phi}}$  has the matrix beta distribution the rows of  $\tilde{\underline{\Phi}}$  are mutually independent random vectors. The decision-maker may, however, wish to use a prior distribution which admits non-zero correlation between the rows of  $\tilde{\underline{\Phi}}$ . Such a distribution may be constructed with the aid of equation (2.3.3), but at the expense of complicating the formulas for the moments. We illustrate this construction for a  $2 \times 2$  random matrix,

$$\tilde{\underline{p}} = \begin{bmatrix} \tilde{p} & 1-\tilde{p} \\ \tilde{q} & 1-\tilde{q} \end{bmatrix}. \quad (6.4.1)$$

A theorem relating to the general  $K \times N$  case is first given.

**Theorem 6.4.1** Let  $h(\underline{\Phi} | \underline{M}, \omega)$  be the probability density function defined by equation (2.3.3) and let  $\mathcal{H}_G$  be the corresponding extended natural conjugate family of distributions. Let  $C(\underline{M}, \omega)$  be the normalising constant defined by equation (2.3.2) and let  $T_{ij}^k(\underline{M})$  be the matrix with the element  $m_{ij}^k$  increased by unity. Then the means, variances, and covariances of  $h(\underline{\Phi} | \underline{M}, \omega)$  are given by

$$E[\tilde{p}_{ij}^k | \underline{M}, \omega] = \tilde{p}_{ij}^k(\underline{M}, \omega) = \frac{C(\underline{M}, \omega)}{C(T_{ij}^k(\underline{M}), \omega)} \quad \begin{array}{l} i=1, \dots, K \\ j=1, \dots, N \end{array} \quad (6.4.2)$$

$$\text{cov}[\tilde{p}_{ij}^{\alpha}, \tilde{p}_{kn}^{\beta} | \underline{M}, \omega] = \tilde{p}_{ij}^{\alpha}(\underline{M}, \omega) [ \tilde{p}_{kn}^{\beta}(T_{ij}^{\alpha}(\underline{M}), \omega) - \tilde{p}_{kn}^{\beta}(\underline{M}, \omega) ]. \quad (6.4.3)$$

$\alpha=1, \dots, K_1$   
 $\beta=1, \dots, K_2$   
 $i, j, k, n=1, \dots, N$

**Proof.** Using equations (2.3.3) and (2.3.2)

$$\begin{aligned} \tilde{p}_{ij}^k(\underline{M}, \omega) &= C(\underline{M}, \omega) \int_{\Omega_{K,N}} p_{ij}^k g(\underline{\Phi} | \omega) \prod_{\alpha=1}^N \prod_{\beta=1}^{K_2} \prod_{v=1}^{K_1} (p_{\alpha\beta}^v)^{m_{ij}^v - 1} d\underline{\Phi} \\ &= \frac{C(\underline{M}, \omega)}{C(T_{ij}^k(\underline{M}), \omega)}, \end{aligned} \quad (6.4.4)$$



which is equation (6.4.2). Theorem 2.3.1 shows that  $\mathcal{H}_g$  is closed under consecutive sampling. Hence, Lemma 2.3.2 is applicable and yields

$$\begin{aligned} E[\tilde{P}_{ij}^{\alpha} \tilde{P}_{km}^{\beta} | \underline{M}, \omega] &= \int_{\mathcal{S}_{k,n}} p_{ij}^{\alpha} P_{km}^{\beta} h(\underline{P} | \underline{M}, \omega) d\underline{P} \\ &= \tilde{p}_{ij}^{\alpha} (\underline{M}, \omega) \tilde{P}_{km}^{\beta} (\underline{P}_{ij}^{\alpha} (\underline{M}), \omega), \end{aligned} \quad (6.4.5)$$

from which (6.4.3) follows. Q.E.D.

Let  $\underline{P}$  be given by (6.4.1) and let

$$\underline{z} = \begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix}, \quad (6.4.6)$$

where  $z_i > 0$  ( $i=1, \dots, 4$ ). Let

$$g(\underline{p}) = (p-q)^2. \quad (6.4.7)$$

Then (2.3.3) becomes

$$h(\underline{p} | \underline{M}) = C(\underline{M})(p-q)^2 p^{m_1-1} q^{m_3-1} (1-p)^{m_2-1} (1-q)^{m_4-1}. \quad \underline{p} \in \mathcal{S}_2 \quad (6.4.8)$$

Evaluating the normalizing constant by means of (2.3.2), we find that

$$C(\underline{M}) = [B(m_1+2, m_2)B(m_3, m_4) - 2B(m_1+1, m_2)B(m_3+1, m_4) + B(m_1, m_2)B(m_3+2, m_4)]^{-1}. \quad (6.4.9)$$

Equation (6.4.2) then yields the means,

$$\begin{aligned} \hat{p}(\underline{M}) &= \frac{B(m_1+2, m_2)B(m_3, m_4) - 2B(m_1+1, m_2)B(m_3+1, m_4) + B(m_1, m_2)B(m_3+2, m_4)}{B(m_1+2, m_2)B(m_3, m_4) - 2B(m_1+1, m_2)B(m_3+1, m_4) + B(m_1, m_2)B(m_3+2, m_4)} \\ &\quad (6.4.10) \end{aligned}$$

$$\begin{aligned} \hat{q}(\underline{M}) &= \frac{B(m_1+2, m_2)B(m_3+1, m_4) - 2B(m_1+1, m_2)B(m_3+2, m_4) + B(m_1, m_2)B(m_3+3, m_4)}{B(m_1+2, m_2)B(m_3, m_4) - 2B(m_1+1, m_2)B(m_3+1, m_4) + B(m_1, m_2)B(m_3+2, m_4)} \\ &\quad (6.4.11) \end{aligned}$$

From (6.4.3) we obtain the covariance



$$\text{cov} [\tilde{P}, \tilde{Q}]_{\underline{y}} =$$

$$\tilde{P}(\underline{y}) \left[ \frac{B(m_1+3, m_2)B(m_3+1, m_4) - 2B(m_1+2, m_2)B(m_3+2, m_4) + B(m_1+1, m_2)B(m_3+3, m_4)}{B(m_1+3, m_2)B(m_3, m_4) - 2B(m_1+2, m_2)B(m_3+1, m_4) + B(m_1+1, m_2)B(m_3+2, m_4)} \right] = \tilde{Q}(\underline{y}) \quad (6.4.12)$$

and it is seen that there is non-zero correlation between the rows of  $\tilde{P}$ .

Let  $\underline{\pi}_{\underline{z}}(P) = (\pi_1(P), \dots, \pi_N(P))$  be the steady state probability vector corresponding to the  $N \times N$  stochastic matrix  $P$  and let  $\underline{v} = (v_1, \dots, v_N)$  be a vector of nonnegative integers. An extended natural conjugate distribution for  $\tilde{P}$  which is required for the analysis to be carried out in Chapter 7 is formed by letting

$$g(\underline{P} | \underline{y}) = \prod_{i=1}^N (\pi_i(P))^{v_i} \quad \text{for } \underline{P} \in \mathcal{S}_N^+ \\ = 0, \quad \text{otherwise} \quad (6.4.13)$$

Let  $\underline{M} = [m_{ij}]$  be an  $N \times N$  matrix of positive elements and let  $\underline{y}_i$  denote the  $i$ th row of  $\underline{M}$  ( $i=1, \dots, N$ ). Then the  $N \times N$  random stochastic matrix  $\tilde{P}$  is said to have the matrix beta-1 distribution with parameter  $(\underline{M}, \underline{y})$  if  $\tilde{P}$  has the joint probability density function

$$f_{M|\underline{y}}^{(N)}(P | \underline{M}, \underline{y}) = W(\underline{M}, \underline{y}) \prod_{i=1}^N \prod_{j=1}^N B_N(m_{ij}) (\pi_i(P))^{v_i} p_{ij}^{m_{ij}-1} \quad \text{for } P \in \mathcal{S}_N^+ \\ = 0, \quad \text{otherwise} \quad (6.4.14)$$

where  $B_N(m_{ij})$  is defined by equation (6.2.4). The normalizing constant  $W(\underline{M}, \underline{y})$  is the reciprocal of  $E[\prod_{i=1}^N (\pi_i(P))^{v_i}]$  when  $\tilde{P}$  has the matrix beta distribution with parameter  $\underline{M}$ ,

$$1/W(\underline{M}, \underline{y}) = \int_{\mathcal{S}_N^+} \prod_{i=1}^N (\pi_i(P))^{v_i} f_{M|\underline{y}}^{(N)}(P | \underline{M}, \underline{y}) dP. \quad (6.4.15)$$

$W(\underline{M}, \underline{y})$  can be computed using the methods of Section 4.2, but this requires



lengthy calculations.

Using Lemma 4.2.1, it is easily seen that

$$\int_{\mathcal{S}_N} f_{M \beta 1}^{(N)} (\underline{P} \mid M, \underline{v}) d\underline{P} = 1. \quad (6.4.16)$$

The first two moments of the distribution are obtained from equations (6.4.2) and (6.4.3),

$$E [\tilde{p}_{1j} \mid M, \underline{v}] = \frac{m_{1j}}{M_1} \cdot \frac{W(M, \underline{v})}{W(T_{1j}^{(N)}, \underline{v})} \quad 1, j=1, \dots, N \quad (6.4.17)$$

$$E [\tilde{p}_{\alpha\beta} \tilde{p}_{\gamma\delta} \mid M, \underline{v}] = \frac{m_{\alpha\beta} m_{\gamma\delta}}{M_\alpha M_\gamma} \cdot \frac{W(M, \underline{v})}{W(T_{\gamma\delta}^{(T_{\alpha\beta}^{(N)})}, \underline{v})} \quad \alpha, \beta, \gamma, \delta = 1, \dots, N \quad (6.4.18)$$

where  $M_1 = \sum_{j=1}^N m_{1j}$  and  $T_{1j}^{(N)}$  is the matrix  $\underline{M}$  with its  $(1, j)$ th element increased by unity.

Due to the complex calculations required to obtain the normalizing constant  $W(\underline{M}, \underline{v})$ , the matrix beta-1 distribution is presently of limited usefulness. This distribution is, however, of some importance since it is the natural conjugate distribution for one of the data-generating processes to be considered in Chapter 7.

## 6.5 The Beta-Whittle Distribution.

The beta-Whittle distribution is defined to be the unconditional distribution of the transition count  $\tilde{\underline{F}}$  of a Markov chain with transition probability matrix  $\tilde{\underline{P}}$  which is drawn from a matrix beta distribution. The beta-Whittle-2 distribution is defined in an analogous fashion. In this section explicit probability mass functions are derived for these distributions and their moments are discussed.



6.5.1 The Beta-Whittle Distribution. For fixed  $u$  and  $v$  ( $u, v = 1, \dots, N$ ) and fixed  $n$  ( $n = 1, 2, 3, \dots$ ), let

$$\Phi_N(u, v, n) = \left\{ \underline{f} \mid f_{ij} \in I; \sum_{i=1}^N \sum_{j=1}^N f_{ij} = n; f_{1i} - f_{i1} = \delta_{iu} - \delta_{iv} \right. \\ \left. (i=1, \dots, N) \right\} \quad (6.5.1)$$

and let

$$\Phi_N(u, n) = \bigcup_{v=1}^N \Phi_N(u, v, n). \quad \begin{matrix} u=1, \dots, N \\ v=1, 2, 3, \dots \end{matrix} \quad (6.5.2)$$

$\Phi_N(u, n)$  is the set of all possible transition counts,  $\underline{F}$ , which can arise from a sample of  $n$  consecutive transitions in a Markov chain with initial state  $u$  and a positive transition probability matrix.

The beta-Whittle probability mass function with parameter  $(u, n, P)$  is defined as

$$f_{\beta P}^{(N)}(\underline{F} \mid u, n, P) = \int_{\Omega_N} \underline{\xi}_N^{(N)}(\underline{F} \mid u, n, P) f_{NP}^{(N, N)}(\underline{P} \mid \underline{M}) d\underline{P} \quad \underline{F} \in \Phi_N(u, n) \\ = 0, \quad \text{elsewhere} \quad (6.5.3)$$

where  $u=1, \dots, N$ ,  $n=1, 2, 3, \dots$ , and  $\underline{M} = [m_{ij}]$  is an  $N \times N$  matrix such that  $m_{ij} > 0$  ( $i, j=1, \dots, N$ ).

When  $\underline{F}$  has the beta-Whittle distribution with parameter  $(u, n, P)$  it is clear that  $\underline{F}$  must have the range set  $\Phi_N(u, n)$ , since the set of stochastic matrices which have one or more elements equal to zero is a set of measure zero relative to the matrix beta distribution.

It is seen from (6.5.3) that  $f_{\beta P}^{(N)}(\underline{F} \mid u, n, P) \geq 0$ . By comparing (6.5.1) and (6.1.5), it is seen that  $\Phi_N(u, n) = \Phi_N(u, n, P)$ , provided  $P$  is a positive matrix. Since the set of nonpositive matrices  $\underline{P}$  is a set of measure zero and since  $\Phi_N(u, n)$  is a finite set, we have



$$\sum_{\substack{F \in \Phi_N(u, n) \\ F \in \Phi_N(u, n, H)}} f_{BW}^{(N)}(F | u, n, H) = \int_{\mathcal{B}_N} \sum_{\substack{H \in \Phi_N(u, n, P) \\ F \in \Phi_N(u, n, H)}} \xi_W^{(N)}(F | u, n, P) f_{MP}^{(N, H)}(P | H) dP$$

$$= 1. \quad (6.5.4)$$

Thus, the beta-Whittle mass function is a proper probability mass function.

**Theorem 6.5.1** The beta-Whittle mass function with parameter  $(u, n, H)$  is given by

$$f_{BW}^{(N)}(F | u, n, H) = \begin{cases} \frac{\prod_{i=1}^N f_{i.} B(f_{i.}, m_{i.})}{\prod_{i=1}^N \prod_{j=1}^N f_{ij} B(f_{ij}, m_{ij})} & F \in \Phi_N(u, n) \\ 0, & \text{elsewhere} \end{cases} \quad (6.5.5)$$

where  $m_{i.} = \sum_{j=1}^N m_{ij}$ ,  $B(x, y)$  is the beta function, and  $v$  is the unique solution of the equations

$$f_{i.} = f_{.i} = \delta_{iu} = \delta_{iv}, \quad i=1, \dots, N$$

**Proof.** Letting  $\underline{m}_i$  denote the  $i$ th row of  $H$ ,

$$f_{BW}^{(N)}(F | u, n, H) = \frac{\prod_{i=1}^N f_{i.}!}{\prod_{i=1}^N \prod_{j=1}^N f_{ij}!} \frac{\prod_{i=1}^N B_N(\underline{m}_i)}{\mathcal{B}_N} \int \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij} + m_{ij}} e^{-1} dP. \quad (6.5.6)$$

The integrand is the kernel of a matrix beta density function with parameter  $H + F$ , hence, using (6.2.4),

$$f_{BW}^{(N)}(F | u, n, H) = \frac{\prod_{i=1}^N \left[ \frac{f_{i.} \Gamma(f_{i.}) \Gamma(m_{i.})}{\Gamma(f_{i.} + m_{i.})} \right]}{\prod_{i=1}^N \prod_{j=1}^N \left[ \frac{\Gamma(f_{ij} + m_{ij})}{f_{ij} \Gamma(f_{ij}) \Gamma(m_{ij})} \right]}$$



$$= \frac{\prod_{i=1}^N f_{1i} \cdot B(f_{1i}, m_{1i})}{\prod_{i=1}^N \prod_{j=1}^N f_{1j} B(f_{1j}, m_{1j})}. \quad (6.5.7)$$

Q.E.D.

The moments of the beta-Whittle distribution are somewhat complicated to compute. Referring to equations (6.1.14) and (6.1.27), if  $\tilde{f}_{ij} = [\tilde{f}_{ij}]$  has the beta-Whittle distribution with parameter  $(u, n, N)$ , then

$$\mathbb{E}[\tilde{f}_{ij}] = E_p E_F | P_{ij} [\tilde{f}_{ij}] = \sum_{k=0}^{n-1} E_p \tilde{p}_{ij}^{(k)} \tilde{p}_{ij}, \quad i, j = 1, \dots, N \quad (6.5.8)$$

and, similarly,

$$\mathbb{E}[\tilde{f}_{\alpha\beta} \tilde{f}_{\gamma\delta}] =$$

$$\delta_{\alpha\gamma} \delta_{\beta\delta} \mathbb{E}[\tilde{f}_{\alpha\beta}] + \sum_{k=1}^{n-1} E_p \tilde{p}_{\alpha\beta}^{(n-1-k)} \tilde{p}_{\alpha\beta} \sum_{v=0}^{k-1} \tilde{p}_{\beta v}^{(v)} \tilde{p}_{\gamma\delta} + \tilde{p}_{\alpha\gamma}^{(n-1-k)} \tilde{p}_{\gamma\delta} \sum_{v=0}^{k-1} \tilde{p}_{\beta v}^{(v)} \tilde{p}_{\alpha\beta} \quad \begin{matrix} \alpha, \beta, \gamma, \delta = 1, \dots, N \\ n = 2, 3, \dots \end{matrix} \quad (6.5.9a)$$

$$= \delta_{\alpha\gamma} \delta_{\beta\delta} \mathbb{E}[\tilde{f}_{\alpha\beta}]. \quad \begin{matrix} \alpha, \beta, \gamma, \delta = 1, \dots, N \\ n = 1 \end{matrix} \quad (6.5.9b)$$

In both of these equations  $E_p[\cdot]$  denotes the expectation operator relative to the distribution  $f_{\alpha\beta}^{(N, N)}(P|N)$ . These expectations can be evaluated by repeated application of Lemma 2.3.2 in a manner which should, by now, be familiar, but the calculations, particularly in (6.5.9a), tend to become extensive. Approximations of the sort we have discussed in Chapter 4 can also be made. For small values of the parameter  $n$ , direct calculation of the moments is probably the most convenient way to approach the problem.



**6.5.2 The Beta-Whittle-2 Distribution.** The set of all possible transition counts  $\underline{F}$  which can arise from a sample of  $n$  consecutive transitions in a Markov chain with arbitrary initial state and a positive transition probability matrix is

$$\Phi_N^*(n) = \bigcup_{u=1}^N \Phi_N^*(u, n). \quad n=1, 2, \dots \quad (6.5.10)$$

The  $N \times N$  random matrix  $\underline{\tilde{F}}$  with range set  $\Phi_N^*(n)$  is said to have the standard beta-Whittle-2 distribution with parameter  $(p, n, M)$  if  $\underline{\tilde{F}}$  has the probability mass function

$$f_{\text{BW2}}^{(N)}(\underline{F} | p, n, M) = \int_{\mathcal{P}_N} f_{\text{BW2}}^{(N)}(\underline{F} | p, n, P) f_{\text{MP}}^{(N, N)}(P | M) dP, \quad (6.5.11)$$

where  $p = (p_1, \dots, p_N)$  is a stochastic vector which is functionally independent of  $P$ ,  $i=1, 2, 3, \dots$ , and  $M = [m_{ij}]$  is an  $N \times N$  matrix with  $m_{ij} > 0$  ( $i, j=1, \dots, N$ ). It is readily established that  $f_{\text{BW2}}^{(N)}(\underline{F} | p, n, M) \geq 0$  and that

$$\sum_{\underline{F} \in \Phi_N^*(n)} f_{\text{BW2}}^{(N)}(\underline{F} | p, n, M) = 1.$$

Let

$$\Phi_{N1}^*(n) = \left\{ \underline{F} \mid \underline{F} \in \Phi_N^*(n), \quad f_{\underline{F}}^{(N)} = f_{\underline{F}_{1.}} \quad (i=1, \dots, N) \right\} \quad (6.5.12)$$

and

$$\Phi_{N2}^*(n) = \Phi_N^*(n) - \Phi_{N1}^*(n). \quad (6.5.13)$$

Since

$$f_{\text{BW2}}^{(N)}(\underline{F} | p, n, P) = \sum_{u=1}^N p_u f_u^{(N)}(\underline{F} | u, n, P),$$

it follows from Lemma 6.1.5 and the fact that  $p$  is functionally independent of  $P$  that



$$\begin{aligned}
 f_{\text{PW2}}^{(N)}(\underline{x} | p, n, \underline{y}) &= \sum_{u=1}^N p_u f_{\text{PW}}^{(N)}(\underline{x} | u, n, \underline{y}) \\
 &= \left( \sum_{i=1}^N p_i F_{ii} \right) \frac{\prod_{i=1}^N f_{ii} B(f_{ii}, n_{ii})}{\prod_{i=1}^N \prod_{j=1}^N f_{ij} B(f_{ij}, n_{ij})} \cdot \frac{p_e \phi_{N2}^e(n)}{N!} \\
 &= p_u F_{uu} \frac{\prod_{i=1}^N f_{ii} B(f_{ii}, n_{ii})}{\prod_{i=1}^N \prod_{j=1}^N f_{ij} B(f_{ij}, n_{ij})} \cdot \frac{p_e \phi_{N2}^e(n)}{N!} \\
 &= 0. \quad \text{elsewhere} \quad (6.5.14)
 \end{aligned}$$

In (6.5.14), if  $\frac{p_e \phi_{N2}^e(n)}{N!}$ ,  $(u, v)$  is the unique solution to the equations

$$f_{ii} = f_{ii} = \delta_{ii} = \delta_{ij} \quad i, j = 1, \dots, N$$

An important case in which  $p$  is not functionally independent of  $P$  occurs when  $p = \underline{\pi}(P)$ , the steady-state probability vector corresponding to  $P$ . In this instance we define the nonstandard beta-Whittle-2 distribution with parameter  $(n, \underline{y})$  in terms of the following probability mass function,

$$f_{\text{PW2e}}^{(N)}(\underline{x} | n, \underline{y}) = \int_N f_{\text{W2}}^{(N)}(\underline{x} | \underline{\pi}_n P) f_{\text{IP}}^{(N, N)}(\underline{x} | \underline{y}) d\underline{p}, \quad (6.5.15)$$

where  $n = 1, 2, 3, \dots$  and  $\underline{y} = [y_{ij}]$  is an  $N \times N$  matrix such that  $y_{ij} > 0$  ( $i, j = 1, \dots, N$ ). The vector  $\underline{\pi}$  in the integrand of (6.5.15) is the steady state vector corresponding to  $P$  and is uniquely defined for all  $P$  except a set of measure zero. It is clear that the range set of  $\underline{\pi}$  is  $\phi_{N2}^e(n)$  and that  $f_{\text{PW2e}}^{(N)}(\underline{x} | n, \underline{y})$  is a proper probability mass function.



**Theorem 6.5.2** If

$$\bar{\pi}_{u \frac{N}{M}}(H) = \int_{\Omega} \pi_u(P) f_{W2}^{(H, M)}(P | H) dP, \quad u=1, \dots, N \quad (6.5.16)$$

is the expected value of  $\pi_u(\tilde{F})$ , then the nonstandard beta-Whittle-2 probability mass function with parameter  $(n, M)$  is given by

$$\begin{aligned} f_{W2}^{(H)}(F | n, M) &= \left( \sum_{i=1}^N \bar{\pi}_i(F + \frac{H}{M}) F_{ii}^* \right) \frac{\prod_{i=1}^N f_{1i} B(f_{1i}, n_{1i})}{\prod_{i=1}^N \prod_{j=1}^N f_{ij} B(f_{ij}, n_{ij})}, \quad F \in \phi_{M1}^*(n) \\ &= \bar{\pi}_u(F + \frac{H}{M}) F_{uu}^* \frac{\prod_{i=1}^N f_{1i} B(f_{1i}, n_{1i})}{\prod_{i=1}^N \prod_{j=1}^N f_{ij} B(f_{ij}, n_{ij})}, \quad F \in \phi_{M2}^*(n) \\ &= 0, \quad \text{elsewhere} \end{aligned} \quad (6.5.17)$$

In (6.5.17), if  $F \in \phi_{M2}^*(n)$ ,  $(u, v)$  is the unique solution to the equations

$$f_{1i} - f_{1j} = \delta_{1x} - \delta_{1y}, \quad i=1, \dots, N$$

**Proof.** Since

$$f_{W2}^{(H)}(F | n, P) = \sum_{i=1}^N \pi_i(P) f_W^{(H)}(F | i, n, P),$$

we have

$$f_{W2}^{(H)}(E | n, H) = \sum_{i=1}^N \int_{\Omega} \pi_i(P) f_W^{(H)}(E | i, n, P) f_{W2}^{(H, M)}(P | H) dP. \quad (6.5.18)$$

The kernel of the integrand of (6.5.18) is

$$\pi_i(P) \prod_{j=1}^N \prod_{k=1}^M P_{jk}^{f_{jk} + n_{jk} - 1}.$$

Thus, proceeding as in the proof of Theorem 6.5.1,



$$f_{\text{BW2}^n}(F | n, \underline{\lambda}) = \sum_{i=1}^N f_{\text{BW}}^{(N)}(F | i, n, \underline{\lambda}) \int_{\underline{\lambda}} \pi_i(p) f_{\text{BW}}^{(N,n)}(p | F + \underline{\lambda}) dp. \quad (6.5.19)$$

Equation (6.5.17) follows from (6.5.19) and Lemma 6.1.5. Q.E.D.

The moments of the standard beta-Whittle-2 distribution can be obtained from the moments of the beta-Whittle distribution by using the relation

$$f_{\text{BW2}^n}^{(N)}(F | p, n, P) = \sum_{i=1}^N p_i f_{\text{BW}}^{(N)}(F | u_i, n, \underline{\lambda}). \quad (6.5.20)$$

The moments of the nonstandard beta-Whittle-2 distribution are given by the following theorem.

**Theorem 6.5.3** Let  $E[\cdot]$  denote the expectation operator relative to the nonstandard beta-Whittle-2 distribution with parameter  $(n, \underline{\lambda})$  and  $E_p[\cdot]$  denote the expectation operator relative to the matrix beta distribution with parameter  $\underline{\lambda}$ . Then

$$E[\tilde{F}_{ij}] = n E_p[\tilde{\pi}_i \tilde{p}_{ij}], \quad i, j = 1, \dots, N \quad (6.5.21)$$

and

$$E[\tilde{F}_{\alpha\beta} \tilde{F}_{\gamma\delta}] = \delta_{\alpha\gamma} \delta_{\beta\delta} E[\tilde{F}_{\alpha\beta}] + \sum_{k=0}^{n-1} (n-1-k) E_p[\tilde{p}_{\alpha\beta} \tilde{p}_{\gamma\delta} (\tilde{\pi}_{\alpha} \tilde{\pi}_{\beta}^{(k)} + \tilde{\pi}_{\gamma} \tilde{\pi}_{\delta}^{(k)})]$$

$$\alpha, \beta, \gamma, \delta = 1, \dots, N \quad (6.5.22a)$$

$n = 2, 3, \dots$

$$= \delta_{\alpha\gamma} \delta_{\beta\delta} E[\tilde{F}_{\alpha\beta}]. \quad \alpha, \beta, \gamma, \delta = 1, \dots, N \quad (6.5.22b)$$

**Proof.** The theorem follows immediately from equations (6.1.53) and (6.1.54), together with the relation

$$E[g(\tilde{F})] = E_p E_{\underline{\lambda}}[g(\tilde{F})],$$

where  $g(F)$  is any function of  $F$  for which the expectation exists and  $E_{\underline{\lambda}}[\cdot]$  is the expectation operator relative to the Whittle-2 distribution with parameter  $(\underline{\lambda}, n, P)$ . Q.E.D.



## CHAPTER 7

### FIXED SAMPLE SIZE ANALYSIS

In Chapters 3-5 we examined some sequential sampling problems in a Markov chain with alternatives. We now consider the prior-posterior and preposterior analysis of a Markov chain governed by a fixed, but unknown,  $N \times N$  matrix of transition probabilities,  $\tilde{P}$ , from which a fixed number of consecutive observations is drawn. In Section 7.1 this analysis is carried out under the assumption that the initial state is known to the decision-maker before the sample is observed. In Section 7.2 it is assumed that the initial state is unknown and has a distribution which is functionally independent of  $\tilde{P}$ ; in the final section it is assumed that the chain is operating in the steady state and that the initial state is unknown.

#### 7.1 Initial State Known

An  $N$ -state Markov chain can be considered to be a process which generates the sequence of random variables,  $\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_i, \dots$ , where  $\tilde{x}_i \in \{1, \dots, N\}$  is the state of the system immediately after the  $i$ th transition ( $i=1, 2, \dots$ ) and  $\tilde{x}_0$  is the initial state observed before the first transition. This initial state,  $\tilde{x}_0 = \tilde{u}$ , is subject to the distribution  $\underline{p} = (p_1, \dots, p_N)$ , where  $\underline{p}$  is a stochastic vector and  $p_i = P[u = i] (i=1, \dots, N)$ . The transitions of the chain are governed by the  $N \times N$  stochastic matrix  $P = [p_{ij}]$ , where  $p_{ij} = P[\tilde{x}_{n+1} = j | \tilde{x}_n = i]$  ( $i, j = 1, \dots, N; n=0, 1, 2, \dots$ ). It is assumed in this section that the initial state is known to the decision-maker. Thus, in this case,



$$p_i = \delta_{iu}.$$

$$i=1, \dots, N \quad (7.1.1)$$

**7.1.1 Prior-Posterior Analysis.** Let  $x_n = (x_0, \dots, x_n)$  be a sample of  $n$  consecutive transitions in a Markov chain, where  $x_0 = u$  is assumed known to the decision-maker before the sample is obtained. Thus,  $x_n$  is obtained under the consecutive sampling rule. Let  $F = [f_{ij}]$  be the transition count of  $x_n$ . Then the conditional probability, given  $\tilde{P} = P$ , of observing the sample  $x_n$  is

$$p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{n-1} x_n} = \prod_{i=1}^n \prod_{j=1}^N p_{ij}^{f_{ij}}. \quad (7.1.2)$$

If the stopping process is noninformative, then (7.1.2) is the kernel of the likelihood of the sample. It is clear that the statistic  $F$  conveys all the information of the sample and that, if stopping is noninformative,  $F$  is a sufficient statistic.

When the transition probability matrix is regarded as a random matrix  $\tilde{P}$ , the natural conjugate of (7.1.2) is the matrix beta distribution defined by equation (6.3.1) with  $K_j = 1$  ( $i=1, \dots, N$ ),

$$\tilde{P}_{MP}^{(N, N)} (P | M) \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{m_{ij}-1}. \quad (7.1.3)$$

If  $\tilde{P}$  has the matrix beta distribution with parameter  $M^* = [m_{ij}^*]$  and if a sample from the process yields a sufficient statistic  $F$ , Theorem 2.2.1 shows that the posterior distribution of  $\tilde{P}$  is matrix beta with parameter

$$M^* = M^* + F. \quad (7.1.4)$$

**7.1.2 Sampling Distributions and Posterior Analysis.** It is assumed that  $x_0 = u$  is known and that  $n$ , the number of transitions to be observed, is determined before the sample is obtained. Prior to sampling, the



transition count  $\tilde{F}$  is a random matrix and the conditional probability, given  $\tilde{P} = P$ , that the Markov chain will generate a specific sample  $x_n$  which has the transition count  $F$  is given by (7.1.2). Whittle [41] has shown that the number of samples of size  $n$  with  $x_n = u$  which have the transition count  $F$  is given by

$$P_{vu} = \frac{\prod_{i=1}^N f_{i.v}}{\prod_{i=1}^N \prod_{j=1}^N f_{ij}}, \quad (7.1.5)$$

where  $f_{i.v} = \sum_{j=1}^N f_{ij}$  ( $i = 1, \dots, N$ ),  $v$  is the final state of the sample, and  $P_{vu}$  is the  $(v,u)$ th cofactor of the matrix  $\tilde{P}$  defined by equation (6.1.8). Thus, the conditional probability of  $\tilde{F}$  is given by the Whittle probability mass function defined by (6.1.7),

$$P[\tilde{F} | u, n, P] = f_W^{(N)}(\tilde{F} | u, n, P). \quad (7.1.6)$$

If a sample of  $n$  consecutive transitions is obtained from a Markov chain with known initial state  $u$  and if the transition matrix  $\tilde{P}$  has the matrix beta distribution with parameter  $N^*$ , then the unconditional distribution of the transition count  $\tilde{F}$  is

$$D(\tilde{F} | u, n, N^*) = \int_{\mathcal{S}_N} f_W^{(N)}(\tilde{F} | u, n, P) f_{MP}^{(N, N)}(P | N^*) dP. \quad (7.1.7)$$

It is seen from equation (6.5.3) that the unconditional distribution of  $\tilde{F}$  is the beta-Whittle mass function given by (6.5.5).

$$D(\tilde{F} | u, n, N^*) = f_{WMP}^{(N)}(\tilde{F} | u, n, N^*). \quad (7.1.8)$$

If the prior distribution of  $\tilde{P}$  is matrix beta with parameter  $N^*$  and if a sample of size  $n$  yields a sufficient statistic  $F$ , then equations (7.1.4)



and (6.3.7) show that the mean of the posterior distribution is

$$\tilde{\bar{p}}^* = [\tilde{\bar{p}}_{ij}^*] \quad (7.1.9)$$

where, if  $\bar{n}_{1.}^* = \sum_{j=1}^N \bar{n}_{1j}^*$ ,

$$\tilde{\bar{p}}_{ij}^* = \frac{\bar{n}_{ij}^* + f_{ij}}{\bar{n}_{1.}^* + f_{1.}} \quad , \quad i, j = 1, \dots, N \quad (7.1.10)$$

Before observing the sample,  $\tilde{\bar{P}}^*$  is a random matrix which can take one of a finite set of values in the range set  $R(u, n, M^*)$ . Let

$$S(P) = \left\{ \bar{P} \mid \bar{P} \in R(u, n), \bar{P}^* = P \right\} \quad \text{PeR}(u, n, M^*) \quad (7.1.11)$$

be the set of possible transition counts which result in a posterior mean with the value  $\text{PeR}(u, n, M^*)$ . Then, by (7.1.8), the distribution of the posterior mean is given by the following probability mass function,

$$P[\tilde{\bar{P}}^* = P \mid u, n, M^*] = \sum_{\bar{P} \in S(P)} \epsilon_{PM}^{(N)}(\bar{P} \mid u, n, M^*) \quad \text{PeR}(u, n, M^*)$$

$$= 0 \quad \text{elsewhere} \quad (7.1.12)$$

## 7.2 Initial State Unknown.

We now assume that the initial state,  $\tilde{x}_0 = \tilde{u}$ , is unknown to the decision-maker before the sample is observed, but that  $\tilde{u}$  has a probability distribution,  $p = (p_1, \dots, p_N)$ , which is functionally independent of  $\tilde{P}$  and which may or may not be known to the decision-maker. If  $\tilde{p}$  is unknown, it is also assumed that the utility of any terminal decision made after  $\tilde{x}_n$  is observed depends only on  $\tilde{P}$  and not on  $\tilde{p}$ .



7.2.1 Prior-Posterior Analysis. Let  $\underline{x}_n = (x_0, \dots, x_n)$  be a sample of  $n$  consecutive transitions in a Markov chain. Let  $u = x_0$  be the initial state observed and let  $\underline{f} = [f_{ij}]$  be the transition count of the sample. Then the conditional probability, given  $\widetilde{\underline{P}} = \underline{P}$  and  $\widetilde{\underline{p}} = \underline{p}$ , of observing the sample  $\underline{x}_n$  is

$$p_{x_0 x_0 x_1 \dots x_{n-1} x_n} = p_u \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}}. \quad (7.2.1)$$

If the stopping process is noninformative, then, since terminal utilities depend only on  $\widetilde{\underline{P}}$  and not on  $\widetilde{\underline{p}}$ , the kernel of the likelihood of the sample is

$$\prod_{i=1}^N \prod_{j=1}^N p_{ij}^{f_{ij}} \quad (7.2.2)$$

and  $\underline{F}$  is a marginally sufficient statistic.

When the matrix of transition probabilities is treated as a random matrix  $\widetilde{\underline{P}}$ , the natural conjugate of (7.2.2) is the matrix beta distribution defined by (6.3.1) with  $K_i = 1$  ( $i=1, \dots, N$ ). If  $\underline{P}$  has the matrix beta distribution with parameter  $M^0 = [m_{ij}]$  and if a sample from the process yields a marginally sufficient statistic  $\underline{F}$ , then the posterior distribution of  $\underline{P}$  is matrix beta with parameter

$$\underline{M}^P = \underline{M}^0 + \underline{F}. \quad (7.2.3)$$

7.2.2 Sampling Distributions and Posterior Analysis. It is assumed that  $n$ , the number of transitions to be observed, is determined before the sample is obtained. Prior to sampling, the pair  $(\widetilde{u}, \widetilde{F})$  is a random quantity and the conditional probability, given  $\widetilde{\underline{P}} = \underline{P}$  and  $\widetilde{\underline{p}} = \underline{p}$ , that the Markov chain will generate a specific sample  $\underline{x}_n$  with the



statistic  $(u, \underline{F})$  is given by (7.2.1). The number of samples of size  $n$  with initial state  $u$  which have the transition count  $\underline{F}$  is given by (7.1.5). Therefore, the conditional probability of  $(\tilde{u}, \tilde{\underline{F}})$  is given by the Whittle-1 probability mass function defined by equation (6.1.37),

$$P[\tilde{u}, \tilde{\underline{F}} | p, n, P] = f_{W1}^{(N)}(\tilde{u}, \tilde{\underline{F}} | p, n, P). \quad (7.2.4)$$

The conditional distribution of the marginally sufficient statistic  $\tilde{\underline{F}}$  is the Whittle-2 probability mass function given by equation (6.1.45),

$$P[\tilde{\underline{F}} | p, n, P] = f_{W2}^{(N)}(\tilde{\underline{F}} | p, n, P). \quad (7.2.5)$$

If a sample of  $n$  consecutive transitions is obtained from a Markov chain where the distribution of the initial state is known to be  $p$  and where the transition probability matrix  $\underline{P}$  has the matrix beta distribution with parameter  $N^0$ , then, provided  $p$  is functionally independent of  $\underline{P}$ , the unconditional distribution of the transition count  $\tilde{\underline{F}}$  is

$$D(\tilde{\underline{F}} | p, n, N^0) = \int_{S_n} f_{W2}^{(N)}(\tilde{\underline{F}} | p, n, P) f_{NP}^{(N, N^0)}(P | N^0) dP. \quad (7.2.6)$$

Thus, by equation (6.5.11), the unconditional distribution of  $\tilde{\underline{F}}$  is the beta-Whittle-2 probability mass function given by (6.5.14),

$$D(\tilde{\underline{F}} | p, n, N^0) = f_{BW2}^{(N)}(\tilde{\underline{F}} | p, n, N^0). \quad (7.2.7)$$

If  $\underline{p}$  is unknown and has the prior distribution function  $H(p | \Psi)$ , with mean  $\bar{p}(\Psi)$ , then equations (7.2.7) and (6.5.14) show that the unconditional distribution of  $\tilde{\underline{F}}$  is also beta-Whittle-2,

$$\begin{aligned} D(\tilde{\underline{F}} | \Psi, n, N^0) &= \int_{S_{1,N}} f_{BW2}^{(N)}(\tilde{\underline{F}} | p, n, N^0) dH(p | \Psi) \\ &= f_{BW2}^{(N)}(\tilde{\underline{F}} | \bar{p}(\Psi), n, N^0). \end{aligned} \quad (7.2.8)$$



If  $\tilde{P}$  has the matrix beta distribution with prior parameter  $M^*$  and if a sample yields the marginally sufficient statistic  $F$ , the mean of the posterior distribution of  $\tilde{P}$  is given by equations (7.1.9) and (7.1.10). Prior to observing the sample the posterior mean,  $\tilde{P}^*$  is a random matrix with the finite range set  $R^*(n, M^*)$ . Let

$$S^*(P) = \left\{ F \mid F \in \phi_N^{(n)}(n), \tilde{P}^* = P \right\} \quad P \in R^*(n, M^*) \quad (7.2.9)$$

be the set of possible transition counts which result in the posterior mean  $\tilde{P}^* = P \in R^*(n, M^*)$ . Then, from (7.2.7), we find that, if  $p$  is known, the distribution of the posterior mean is given by the following probability mass function,

$$\begin{aligned} P[\tilde{P}^* = P \mid p, n, M^*] &= \sum_{F \in S^*(P)} \frac{\epsilon_{\beta+2}^{(n)}}{\epsilon_{\beta+2}^{(N)}} (F \mid p, n, M^*) \quad P \in R^*(n, M^*) \\ &= 0. \quad \text{elsewhere} \end{aligned} \quad (7.2.10)$$

Similarly, if  $\tilde{p}$  is unknown and has the prior distribution function  $H(p \mid \gamma)$ , the distribution of the posterior mean is

$$\begin{aligned} P[\tilde{P}^* = P \mid \gamma, n, M^*] &= \sum_{F \in S^*(P)} \frac{\epsilon_{\beta+2}^{(n)}}{\epsilon_{\beta+2}^{(N)}} (F \mid \tilde{p}(\gamma), n, M^*) \quad P \in R^*(n, M^*) \\ &= 0. \quad \text{elsewhere} \end{aligned} \quad (7.2.11)$$

### 7.3 System Operating in the Steady State.

When the Markov chain is operating in the steady state and the initial state,  $\tilde{u}$ , is unknown, the distribution of  $\tilde{u}$  is  $\pi(\tilde{P}) = (\pi_1(\tilde{P}), \dots, \pi_N(\tilde{P}))$ , the steady-state probability vector associated with the transition matrix  $\tilde{P}$ . In this case, observation of  $\tilde{u}$  provides information about  $\tilde{P}$ .

**7.3.1 Prior-Posterior Analysis.** Let  $x_n = (x_0, \dots, x_n)$  be a sample of  $n$  consecutive transitions in a Markov chain which is operating in the



steady state. If  $u = x_0$  is the initial state and  $\underline{y} = [y_{ij}]$  is the transition count of the sample, the conditional probability, given that  $\tilde{\underline{P}} = \underline{P}$ , of observing the sample  $\underline{x}_n$  is

$$\pi_{x_0}(P) p_{x_0 x_1} \cdots p_{x_{n-1} x_n} = \pi_u(P) \prod_{i=1}^N \prod_{j=1}^N p_{ij}^{y_{ij}}. \quad (7.3.1)$$

When stopping is noninformative, equation(7.3.1) is the kernel of the likelihood of the sample and the ordered pair  $(u, \underline{y})$  is a sufficient statistic.

When  $\tilde{\underline{P}}$ , the matrix of transition probabilities, is regarded as a random matrix the natural conjugate of equation (7.3.1) is the matrix beta-1 distribution defined by (6.4.14),  $f_{MB1}^{(N)}(\underline{P} | \underline{y}, \underline{v})$ . It is easily seen that, if  $\tilde{\underline{P}}$  has the matrix beta-1 distribution with parameter  $(\underline{M}^0, \underline{v}^0)$  and if a sample from the process yields a sufficient statistic  $(u, \underline{y})$ , the posterior distribution of  $\tilde{\underline{P}}$  is matrix beta-1 with parameter  $(\underline{M}^u, \underline{v}^u)$ , where, if  $\underline{e}_u$  is an  $N$ -dimensional row vector with  $u$ th component equal to one and all other components equal to zero,

$$\begin{matrix} \underline{M}^u \\ \underline{v}^u \end{matrix} = \begin{matrix} \underline{M}^0 \\ \underline{v}^0 \end{matrix} + \begin{matrix} \underline{P} \\ \underline{e}_u \end{matrix} \quad (7.3.2a)$$

$$\underline{v}^u = \underline{v}^0 + \underline{e}_u. \quad (7.3.2b)$$

As was noted in Section 6.4, the normalizing constant and the moments of the matrix beta-1 distribution are difficult to compute. This difficulty complicates the task of assigning a specific matrix beta-1 prior distribution to  $\tilde{\underline{P}}$ . Since the matrix beta distribution is also a matrix beta-1 distribution with the parameter  $\underline{v} = (0, \dots, 0) = \underline{0}$ ,

$$f_{MB}^{(N,N)}(\underline{P} | \underline{y}) = f_{MB1}^{(N)}(\underline{P} | \underline{y}, \underline{0}), \quad (7.3.3)$$

it may be convenient for the decision-maker to use a matrix beta prior distribution for  $\tilde{\underline{P}}$  and we shall assume this to be the case in discussing



the preposterior analysis of a Markov chain operating in the steady state.

**7.3.2 Sampling Distributions and Preposterior Analysis.** We assume that  $n$ , the number of transitions to be observed, is fixed in advance of sampling and that the prior distribution of  $\tilde{P}$  is matrix beta. Prior to sampling, the conditional probability, given  $\tilde{P} = P$ , of obtaining a specific sample  $\tilde{x}_n$  with the statistic  $(\tilde{u}, \tilde{F})$  is given by (7.3.1). The number of samples of size  $n$  with initial state  $u$  which have the transition count  $\tilde{F}$  is given by (7.1.5) and, therefore, the conditional probability of the statistic  $(\tilde{u}, \tilde{F})$  is given by the Whittle-1 probability mass function as defined by (6.1.37),

$$P(u, \tilde{F} | n, P) = f_{W1}^{(N)}(u, \tilde{F} | \underline{\Pi}(P), n, P). \quad (7.3.4)$$

The marginal conditional distribution of  $\tilde{u}$  is  $\underline{\Pi}(P)$  and the marginal conditional distribution of  $\tilde{F}$  is the Whittle-2 distribution,  
 $f_{W2}^{(N)}(\tilde{F} | \underline{\Pi}(P), n, P).$

When a sample  $x_n$  is obtained from a Markov chain operating in the steady state where the initial state is unknown and the transition probability matrix  $\tilde{P}$  has the matrix beta distribution with parameter  $M^0$ , the unconditional distribution of the transition count  $\tilde{F}$  is

$$D(\tilde{F} | n, M^0) = \int_{S_n} f_{W2}^{(N)}(\tilde{F} | \underline{\Pi}(P), n, P) f_{M^0}^{(N, N)}(P | M^0) dP. \quad (7.3.5)$$

Therefore, using equation (6.5.15), the unconditional distribution of  $\tilde{F}$  is nonstandard beta-Whittle-2, as given by equation (6.5.17).

$$D(\tilde{F} | n, M^0) = f_{SN2}^{(N)}(\tilde{F} | n, M^0) \quad (7.3.6)$$

It is then easily seen that, if the set  $S^*(P)$  is defined by equation (7.2.9), the prior distribution of the posterior mean is given by the



following probability mass function,

$$P[\tilde{P}^n = p \mid n_0 M^0] = \sum_{\substack{\text{FeS} \\ \in P}} f_{\text{FeS}}^{(n)} (p \mid n_0 M^0) \quad \text{for } p \in \text{FeS} \\ = 0. \quad \text{elsewhere} \quad (7.3.7)$$



CHAPTER 8  
SPECIFIC RESULTS FOR A  
TWO-STATE MARKOV CHAIN

Many of the matters considered in previous chapters are specialized to the case of a two-state Markov chain in this chapter. The  $2 \times 2$  transition probability matrix  $\tilde{P}$  is assumed to have the matrix beta distribution and explicit formulas are found for the means and product moments of the  $n$ -step transition probabilities, the steady-state probabilities, the process gain, and the expected total discounted rewards. The chapter concludes with a result concerning the selection of an optimal terminal policy for a two-state process with a special type of reward structure. Most of the formulas derived here are double infinite series; it appears doubtful that similar expressions can be obtained for chains with more than two states.

### 8.1 Preliminaries.

Let

$$\begin{matrix} P = & \begin{bmatrix} 1-x & x \\ y & 1-y \end{bmatrix} & 0 \leq x, y \leq 1 \end{matrix} \quad (8.1.1)$$

be the transition probability matrix for a two-state Markov chain. The eigenvalues of  $P$  are the roots of the equation

$$|\lambda I - P| = \lambda^2 - (2-x-y)\lambda + (1-x-y) = 0 \quad (8.1.2)$$

and are found to be

$$\lambda_1 = 1 \quad 0 \leq x, y \leq 1 \quad (8.1.3a)$$

$$\lambda_2 = 1-x-y \quad 0 \leq x, y \leq 1 \quad (8.1.3b)$$



The eigenvalues of  $P$  are distinct provided  $x$  and  $y$  are not both equal to zero. When  $\lambda_1 \neq \lambda_2$ , Sylvester's Theorem leads to the spectral decomposition

$$P = \begin{bmatrix} -\frac{y}{x+y} & \frac{x}{x+y} \\ \frac{y}{x+y} & \frac{x}{x+y} \end{bmatrix} + (1-x-y) \begin{bmatrix} \frac{x}{x+y} & \frac{-x}{x+y} \\ \frac{-y}{x+y} & \frac{y}{x+y} \end{bmatrix}. \quad (8.1.4)$$

$x \neq 0$  or  $y \neq 0$

Equation (8.1.4) immediately gives the following expressions for the steady-state vector,

$$\pi(P) = \left[ \frac{y}{x+y}, \frac{x}{x+y} \right], \quad (8.1.5)$$

$x \neq 0$  or  $y \neq 0$

and the  $n$ -step probability matrix,

$$P^{\mu} = \begin{bmatrix} p_{11}^{(\mu)} & p_{12}^{(\mu)} \\ p_{21}^{(\mu)} & p_{22}^{(\mu)} \end{bmatrix} = \begin{bmatrix} \frac{y}{x+y} & \frac{x}{x+y} \\ \frac{y}{x+y} & \frac{x}{x+y} \end{bmatrix} + (1-x-y)^{\mu} \begin{bmatrix} \frac{x}{x+y} & \frac{-x}{x+y} \\ \frac{-y}{x+y} & \frac{y}{x+y} \end{bmatrix}. \quad (8.1.6)$$

$\mu=0, 1, 2, \dots$   
 $x \neq 0$  or  $y \neq 0$

In particular,

$$p_{12}^{(\mu)} = \frac{x}{x+y} [1 - (1-x-y)^{\mu}] = x \frac{1 - (1-x-y)^{\mu}}{1 - (1-x-y)} \\ = x \sum_{k=0}^{\mu-1} (1-x-y)^k, \quad \mu=1, 2, 3, \dots \quad (8.1.7)$$

$x \neq 0$  or  $y \neq 0$

and, similarly,

$$p_{21}^{(\mu)} = y \sum_{k=0}^{\mu-1} (1-x-y)^k. \quad \mu=1, 2, 3, \dots \quad (8.1.8)$$

$x \neq 0$  or  $y \neq 0$

Let the process have the reward matrix

$$R = [r_{ij}] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (8.1.9)$$

where  $r_{ij}$  is the reward earned when the process makes a transition from state  $i$  to state  $j$  ( $-\infty < r_{ij} < \infty$ ).



Let

$$\mathbf{H} = \begin{bmatrix} m & n \\ p & q \end{bmatrix} \quad (8.1.10)$$

and assume that  $\tilde{\mathbf{P}}$  has the matrix beta distribution with parameter  $\mathbf{H}$ ,

$$f_{\text{MB}}^{(2,2)}(\mathbf{x} | \mathbf{H}) = \frac{1}{B(m,n)B(p,q)} x^{m-1}(1-x)^{n-1} y^{p-1}(1-y)^{q-1}. \quad (8.1.11)$$

Thus,  $\tilde{x}$  and  $\tilde{y}$  are independent random variables, each having the univariate beta distribution. It is to be noted that equations (8.1.7) and (8.1.8) are valid for all  $\mathbf{x} \in \mathbb{S}_2$  except a set of measure zero relative to the matrix beta distribution.

## 6.2 Hypergeometric Coefficients.

It will be convenient to use the hypergeometric coefficient,  $(x)_k$ , in the equations of subsequent sections. This coefficient is defined here and some of its properties are derived.

Let  $x$  be any real number and  $k$  any nonnegative integer. The hypergeometric coefficient is defined by

$$(x)_k = x(x+1) \dots (x+k-1), \quad k=1, 2, \dots \quad (8.2.1a)$$

$$= 1. \quad k=0 \quad (8.2.1b)$$

If  $x > 0$  it is clear that

$$(x)_k = \frac{\Gamma(x+k)}{\Gamma(x)} \quad (8.2.2)$$

and in the case  $x = 1$ ,

$$(1)_k = k! \quad (8.2.3)$$

**Lemma 8.2.1** If  $(x)_k$  is the hypergeometric coefficient defined by (8.2.1), then the following relations hold,

$$x(x+1)_k = (x)_k (x+k) \quad (8.2.4)$$



$$x(n+1)_k = (x)_k^{n+1} \quad (8.2.5)$$

$$(x)_{k+1} = (x)_k (n+k) \quad (8.2.6)$$

$$x(n+1)_{k+1} = (x)_k (n+k)(n+k+1) \quad (8.2.7)$$

$$(x)_k (n+k)_v = (x)_{k+v}. \quad (8.2.8)$$

Proof. Equations (8.2.4) and (8.2.5) follow by writing

$$x(n+1)_k = x(n+1) \dots (n+k) = (x)_k (n+k) = (x)_{k+1}. \quad (8.2.9)$$

Equation (8.2.6) follows directly from (8.2.4) and (8.2.5). To obtain (8.2.7) we use (8.2.6) and (8.2.4) to obtain

$$\begin{aligned} x(n+1)_{k+1} &= x(n+1)_k (n+k+1) \\ &= (x)_k (n+k)(n+k+1). \end{aligned} \quad (8.2.10)$$

Equation (8.2.8) follows by direct expansion,

$$\begin{aligned} (x)_k (n+k)_v &= x(n+1) \dots (n+k-1)(n+k) \dots (n+k+v-1) \\ &= (x)_{k+v}. \end{aligned} \quad (8.2.11)$$

G.E.D.

### 8.3 Expected instant Transition Probabilities.

We first consider the expected value of  $\hat{p}_{12}^{(\mu)}$ . Using the binomial theorem to expand the factor  $(1-x-y)^k$  of equation (8.1.7),

$$(1-x-y)^k = \sum_{v=0}^k \binom{k}{v} (-1)^v y^v (1-y)^{k-v}, \quad \begin{matrix} k=0, 1, 2, \dots \\ 0 \leq x, y \leq 1 \end{matrix} \quad (8.3.1)$$

we can write



$$\begin{aligned} E[\tilde{P}_{12}^{(\mu)}] &= \int_0^1 x \sum_{k=0}^{\mu-1} (1-x-y)^k f_{xy}^{(2,2)}(x|y) dP \\ &= \sum_{k=0}^{\mu-1} \sum_{v=0}^k \binom{k}{v} (-1)^v E_y[\tilde{y}^v] E_x[\tilde{x}(1-\tilde{x})^{k-v}]. \end{aligned} \quad (8.3.2)$$

For  $a = 0, 1, 2, \dots$

$$\begin{aligned} E_x[\tilde{x}(1-\tilde{x})^a] &= \frac{1}{B(n,n)} \int_0^1 x^{(n+1)-1} (1-x)^{n+a-1} dx \\ &= \frac{B(n+a, n+1)}{B(n,n)} = \frac{n}{n+n} \frac{(n)_a}{(n+n+1)_a} \end{aligned} \quad (8.3.3)$$

and

$$\begin{aligned} E_y[\tilde{y}^a] &= \frac{1}{B(p,q)} \int_0^1 y^{(p+a)-1} (1-y)^{q-1} dy \\ &= \frac{B(p+a+q)}{B(p,q)} = \frac{(p)_a}{(p+q)_a}. \end{aligned} \quad (8.3.4)$$

Thus we have

$$E[\tilde{P}_{12}^{(\mu)}] = \frac{n}{n+n} \sum_{k=0}^{\mu-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v} (p)_v}{(n+n+1)_{k-v} (p+q)_v}, \quad (8.3.5)$$

$\mu=1, 2, 3, \dots$

The following recurrence relation, which follows immediately from (8.3.5), is of use for computing successive values of  $E[\tilde{P}_{12}^{(\mu)}]$ ,

$$E[\tilde{P}_{12}^{(\mu+1)}] = E[\tilde{P}_{12}^{(\mu)}] + \frac{n}{n+n} \sum_{v=0}^{\mu} \binom{\mu}{v} (-1)^v \frac{(n)_{\mu-v} (p)_v}{(n+n+1)_{\mu-v} (p+q)_v}, \quad (8.3.6)$$

$\mu=1, 2, 3, \dots$

In a similar fashion an expression for  $E[\tilde{P}_{21}^{(\mu)}]$  is easily derived, using (8.1.8).

$$E[\tilde{P}_{21}^{(\mu)}] = \sum_{k=0}^{\mu-1} \sum_{v=0}^k \binom{k}{v} (-1)^v E_y[y^{v+1}] E_x[(1-\tilde{x})^{k-v}]$$



$$= \sum_{k=0}^{\mu-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{\binom{n}{k-v} \binom{p}{v+1}}{\binom{n+v}{k-v} \binom{p+q}{v+1}}. \quad (8.3.7)$$

$\mu = 1, 2, \dots$

For purposes of computation, we have the recurrence relation

$$E[\tilde{p}_{21}^{(\mu+1)}] = E[\tilde{p}_{21}^{(\mu)}] + \sum_{v=0}^{\mu} \binom{\mu}{v} (-1)^v \frac{\binom{n}{\mu-v} \binom{p}{v+1}}{\binom{n+v}{\mu-v} \binom{p+q}{v+1}}. \quad (8.3.8)$$

$\mu = 1, 2, \dots$

The derivation of (8.3.5) and (8.3.7) depended upon the form of the expressions (8.1.7) and (8.1.8). Similar expressions cannot be obtained for  $\tilde{p}_{11}^{(\mu)}$  and  $\tilde{p}_{22}^{(\mu)}$ . Thus, the diagonal elements of the mean n-step transition probability matrix must be computed from the relations

$$E[\tilde{p}_{11}^{(\mu)}] = 1 - E[\tilde{p}_{12}^{(\mu)}] \quad (8.3.9a)$$

and

$$E[\tilde{p}_{22}^{(\mu)}] = 1 - E[\tilde{p}_{21}^{(\mu)}]. \quad (8.3.9b)$$

We now verify that  $E[\tilde{p}_{12}^{(\mu)}]$  satisfies the recursive equation (4.1.2). That is, we shall show that  $\tilde{p}_{12}^{(\mu)}(M) = E[\tilde{p}_{12}^{(\mu)}]$  satisfies

$$\tilde{p}_{12}^{(\mu+1)}(M) = \sum_{k=1}^2 \tilde{p}_{1k}^{(\mu)}(T_{k2}(M)) \tilde{p}_{k2}^{(\mu)}(M), \quad (8.3.10)$$

where  $\tilde{p}_{ij}(M)$  is the expected value of  $\tilde{p}_{ij}$  when  $\tilde{P}$  has the distribution  $\xi_{np}^{(2,2)}(P | M)$  and where  $T_{ij}(M)$  is the parameter matrix  $M$  with its  $(i,j)$ th element increased by unity.

Since

$$\tilde{p}_{12}(M) = \frac{1}{n+m} \quad (8.3.11a)$$

$$\tilde{p}_{22}(M) = \frac{1}{p+q}, \quad (8.3.11b)$$

the right side of (8.3.10) is



$$\begin{aligned} & \frac{n}{m+n} \left[ 1 - \frac{n+1}{m+n+1} \sum_{k=0}^{n-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+2)_{k-v}(p+q)_v} \right] \\ & + \frac{q}{p+q} \frac{n}{m+n} \sum_{k=0}^{n-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q+1)_v}. \quad (8.3.12) \end{aligned}$$

Using (8.2.4), equation (8.3.11) can be written

$$\frac{n}{m+n} \left[ 1 + \sum_{k=0}^{n-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v} \left[ \frac{q}{p+q+v} - \frac{n+1}{m+n+1+k-v} \right] \right]. \quad (8.3.13)$$

Since

$$\frac{q}{p+q+v} = \frac{n+1}{m+n+1+k-v} = \frac{m+k-v}{m+n+1+k-v} = \frac{p+v}{p+q+v} \quad (8.3.14)$$

(8.3.13) becomes, upon applying (8.2.6),

$$\begin{aligned} & \frac{n}{m+n} \left[ 1 + \sum_{k=0}^{n-1} \sum_{v=0}^k \binom{k}{v} (-1)^{v+1} \frac{(n)_{k-v}(p)_{v+1}}{(m+n+1)_{k-v}(p+q)_{v+1}} \right. \\ & \left. + \sum_{k=0}^{n-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k+1-v}(p)_v}{(m+n+1)_{k+1-v}(p+q)_v} \right]. \quad (8.3.15) \end{aligned}$$

Letting  $j = v+1$  in the first sum and noting that  $\binom{k}{v-1} + \binom{k}{v} = \binom{k+1}{v}$ , we obtain

$$\begin{aligned} & \frac{n}{m+n} \left[ 1 + \sum_{k=0}^{n-1} \left[ \frac{(n)_{k+1}}{(m+n+1)_{k+1}} + \sum_{v=1}^k \binom{k+1}{v} (-1)^v \frac{(n)_{k+1-v}(p)_v}{(m+n+1)_{k+1-v}(p+q)_v} \right. \right. \\ & \left. \left. + (-1)^{k+1} \frac{(p)_{k+1}}{(p+q)_{k+1}} \right] \right]. \quad (8.3.16) \end{aligned}$$

which, upon letting  $j = k+1$  and collecting terms, is  $\tilde{P}_{12}^{(\mu+1)}(\frac{N}{2})$ , as required. A similar derivation shows that  $\tilde{P}_{21}^{(\mu)}(\frac{N}{2})$ , as defined by (8.3.7),



satisfies (4.1.2).

#### 8.4 Expected Value of $\tilde{P}_{12}^{(\mu)} \tilde{P}_{21}^{(\mu)}$ .

Using (8.1.7), we have, for fixed  $P_x$ ,

$$(\tilde{P}_{12}^{(\mu)})^2 = x^2 \sum_{j=1}^{\mu-1} \sum_{k=0}^{\mu-1} (1-x-y)^{j+k} \quad (8.4.1)$$

and, therefore,

$$E[(\tilde{P}_{12}^{(\mu)})^2] = \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v E[y^v] E[x^2(1-x)^{j+k-v}]. \quad (8.4.2)$$

Since

$$E[x^2(1-x)^a] = \frac{B(m+a, m+2)}{B(m, m)} = \frac{(n)_2}{(m+n)_2} \frac{(n)_a}{(m+n+2)_a} \quad (8.4.3)$$

$$a=0, 1, 2, \dots$$

we have

$$E[(\tilde{P}_{12}^{(\mu)})^2] = \frac{(n)_2}{(m+n)_2} \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k-v} (p)_v}{\binom{m+n+2}{j+k-v} (p+q)_v}. \quad (8.4.4)$$

$$\mu=1, 2, \dots$$

Similarly, by equation (8.1.8),

$$(\tilde{P}_{21}^{(\mu)})^2 = y^2 \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} (1-x-y)^{j+k}, \quad (8.4.5)$$

and

$$E[(\tilde{P}_{21}^{(\mu)})^2] = \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k+v} (p)_{v+2}}{\binom{m+n}{j+k+v} (p+q)_{v+2}}. \quad (8.4.6)$$

$$\mu=1, 2, \dots$$

Finally, since



$$P_{12}^{(\mu)} P_{21}^{(\mu)} = \frac{n}{m+n} \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} (1-\alpha y)^{j+k}, \quad (8.4.7)$$

we have

$$E[P_{12}^{(\mu)} P_{21}^{(\mu)}] = \frac{n}{m+n} \sum_{j=0}^{\mu-1} \sum_{k=0}^{\mu-1} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k-v} \binom{p}{v+1}}{\binom{m+n+1}{j+k-v} \binom{p+q}{v+1}}. \\ \mu=1, 2, \dots \quad (8.4.8)$$

The same method can be used to derive more general product moments of the form  $E[\tilde{P}_{\alpha\beta}^{(\mu)} \tilde{P}_{\gamma\delta}^{(\nu)}]$ .

### 8.5 Steady-State Probabilities.

We now obtain expressions for the means and product moments of  $\pi(\tilde{P}) = (\tilde{\pi}_1, \tilde{\pi}_2)$ . Silver [38], treating the special case where  $y$  is known and  $\tilde{x}$  has the beta distribution, has shown that  $E[\tilde{\pi}_1]$  is a Gaussian hypergeometric function.\*

By Theorem 4.2.5,  $\lim_{\mu \rightarrow \infty} E[\tilde{P}_{12}^{(\mu)}] = E[\tilde{\pi}_1]$  and using equations (8.3.7) and (8.3.5), we immediately have

$$E[\tilde{\pi}_1] = \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{\binom{n}{k-v} \binom{p}{v+1}}{\binom{m+n}{k-v} \binom{p+q}{v+1}} \quad (8.5.1)$$

$$E[\tilde{\pi}_2] = \frac{n}{m+n} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{\binom{n}{k-v} \binom{p}{v}}{\binom{m+n+1}{k-v} \binom{p+q}{v}}. \quad (8.5.2)$$

Theorem 4.2.5 implies that the series (8.5.1) and (8.5.2) both converge. We shall show that they converge conditionally. Neglecting the constant multiplier  $\frac{n}{m+n}$ , and noting that  $\binom{k}{v} = \binom{k}{k-v}$ , the series of absolute values corresponding to (8.5.2) is

---

\* Erdélyi [17], ch. 2.



$$\begin{aligned}
 & \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} \frac{(n)_{k-v} (p)_v}{(m+n+1)_{k-v} (p+q)_v} = \sum_{v=0}^{\infty} \sum_{k=0}^{\infty} \binom{vk}{k} \frac{(n)_k (p)_v}{(m+n+1)_k (p+q)_v} \\
 & = \sum_{v=0}^{\infty} \sum_{k=0}^{\infty} \frac{(1)_{vk} (n)_k (p)_v}{(m+n+1)_k (p+q)_v} = \frac{1}{k!} \frac{1}{v!} \\
 & = F_2(1, m, p, m+n+1, p+q; 1, 1), \tag{8.5.3}
 \end{aligned}$$

where  $F_2(\alpha, \beta, \beta'; \gamma, \gamma'; x, y)$  is Appell's second hypergeometric function of two arguments [2]. Since  $F_2(\alpha, \beta, \beta'; \gamma, \gamma'; x, y)$  diverges whenever  $|x| + |y| > 1$ , the series (8.5.3) diverges and, therefore, the series (8.5.2) converges conditionally. A similar proof establishes the conditional convergence of (8.5.1).

It is easily verified that  $E[\tilde{\pi}_1]$  and  $E[\tilde{\pi}_2]$  satisfy equation (4.2.40a). Let  $\tilde{\pi}_j(\underline{y}) = E[\tilde{\pi}_j]$  ( $j=1, 2$ ). Then it must be shown that  $\tilde{\pi}_j(\underline{y})$  satisfies

$$\tilde{\pi}_j(\underline{y}) = \sum_{k=1}^2 \tilde{\pi}_k(T_k \underline{y}) \tilde{R}_{kj}(\underline{y}). \quad j=1, 2 \tag{8.5.4}$$

We shall consider the case  $j=2$ ; the proof for  $j=1$  is similar.

For  $j=2$ , the right side of (8.5.4) is

$$\begin{aligned}
 & \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} (-1)^v \left[ \frac{n}{m+n} \frac{(n)_{k-v} (p)_{v+1}}{(m+n+1)_{k-v} (p+q)_{v+1}} + \frac{n}{m+n} \frac{q}{p+q} \frac{(n)_{k-v} (p)_v}{(m+n+1)_{k-v} (p+q+1)_v} \right] \\
 & = \frac{n}{m+n} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}}{(m+n+1)_{k-v}} \left[ \frac{(p)_{v+2}}{(p+q)_{v+1}} + \frac{q(p)_v}{(p+q)_{v+1}} \right]. \tag{8.5.5}
 \end{aligned}$$

By using (8.2.6) we see that

$$\frac{(p)_{v+1} + q(p)_v}{(p+q)_{v+1}} = \frac{(p)_v (p+qv)}{(p+q)_v (p+qv)} = \frac{(p)_v}{(p+q)_v}, \tag{8.5.6}$$

and, therefore, that (8.5.5) is equal to  $\tilde{\pi}_2(\underline{y})$ .



By Theorem 4.2.8,  $\lim_{\mu \rightarrow \infty} E[\tilde{\pi}_{\alpha \beta}^{(\mu)} \tilde{\pi}_{\gamma \delta}^{(\mu)}] = E[\tilde{\pi}_{\alpha \beta} \tilde{\pi}_{\gamma \delta}]$  and we obtain the following equations from (8.4.6), (8.4.4), and (8.4.8),

$$E[\tilde{\pi}_1^2] = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k-v} \binom{p}{v+2}}{\binom{m+n}{j+k-v} \binom{p+q}{v+2}} \quad (8.5.7)$$

$$E[\tilde{\pi}_2^2] = \frac{\binom{n}{2}}{\binom{m+n}{2}} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k-v} \binom{p}{v}}{\binom{m+n+2}{j+k-v} \binom{p+q}{v}} \quad (8.5.8)$$

$$E[\tilde{\pi}_2 \tilde{\pi}_1] = \frac{\binom{n}{2}}{\binom{m+n}{2}} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{j+k} \binom{j+k}{v} (-1)^v \frac{\binom{n}{j+k-v} \binom{p}{v+1}}{\binom{m+n+1}{j+k-v} \binom{p+q}{v+1}} \quad (8.5.9)$$

The series (8.5.7) ~ (8.5.9) are conditionally convergent. We illustrate the proof for equation (8.5.8). By Theorem 4.2.8, the double infinite series (8.5.8) converges. Neglecting the constant multiplier

$\frac{\binom{n}{2}}{\binom{m+n}{2}}$ , the corresponding series of absolute values is

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{j+k} \binom{j+k}{v} \frac{\binom{n}{j+k-v} \binom{p}{v}}{\binom{m+n+2}{j+k-v} \binom{p+q}{v}}. \quad (8.5.10)$$

Using (8.5.3) we can write equation (8.5.10) as

$$F_2(1, n, p, m+n+2, p+q; 1, 1) + \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{j+k} \binom{j+k}{v} \frac{\binom{n}{j+k-v} \binom{p}{v}}{\binom{m+n+2}{j+k-v} \binom{p+q}{v}}, \quad (8.5.11)$$

which diverges. Thus (8.5.8) is conditionally convergent. Similar proofs show that (8.5.7) and (8.5.9) also converge conditionally.

It can be verified that  $E[\tilde{\pi}_1 \tilde{\pi}_j]$  satisfies equation (4.2.55a). The algebra is straightforward but tedious and will not be reproduced here.



### 8.6 Process Gain.

The expected gain of the two-state Markov chain considered in this chapter is, by (4.4.3),

$$\bar{G}(H) = \sum_{i=1}^2 \sum_{j=1}^2 \bar{\pi}_i(T_{ij}(H)) \bar{P}_{ij}(H) r_{ij}. \quad (8.6.1)$$

If the reward matrix  $R$  is given by (8.1.9), the expected gain is

$$\begin{aligned} \bar{G}(H) &= \sum_{k=0}^{\infty} \sum_{v=0}^k (v)(-1)^v \left[ a \frac{n}{m+n} \frac{(m)_k (p)_{v+1}}{(m+n+1)_{k-v} (p+q)_{v+1}} \right. \\ &\quad + b \frac{n}{m+n} \frac{(n)_k (p)_{v+1}}{(m+n+1)_{k-v} (p+q)_{v+1}} \\ &\quad \left. + c \frac{n}{m+n} \frac{p (m)_k (p+1)_v}{p+q (m+n+1)_{k-v} (p+q+1)_v} \right] \\ &\quad + d \frac{n}{m+n} \frac{p}{p+q} \frac{(n)_k (p)_v}{(m+n+1)_{k-v} (p+q+1)_v}. \end{aligned} \quad (8.6.2)$$

Applying (8.2.4), (8.2.5) and (8.2.6), we have

$$\bar{G}(H) = \frac{n}{m+n} \sum_{k=0}^{\infty} \sum_{v=0}^k (v)(-1)^v \frac{(m)_k (p)_{v+1}}{(m+n+1)_{k-v} (p+q)_{v+1}} \left[ \frac{a}{n} (m+k-v) + b + c \frac{d}{p+q} \right]. \quad (8.6.3)$$

It is clear that (8.6.3) must converge conditionally.

### 8.7 Total Discounted Reward Vector.

The expected discounted reward over an infinite period when the system starts in state 1 is given by equation (4.3.7) as

$$\bar{V}_1(H) = \sum_{\mu=0}^{\infty} \beta^{\mu} \sum_{j=1}^2 \sum_{k=1}^2 \bar{P}_{1,j}^{(\mu)}(T_{jk}(H)) \bar{P}_{jk}(H) r_{jk}, \quad 1=1,2, \quad 0 \leq \beta < 1 \quad (8.7.1)$$

Let

$$S_{1,j}(H) = \sum_{\mu=1}^{\infty} \beta^{\mu} \bar{P}_{1,j}^{(\mu)}(H). \quad 1, j=1, 2, \quad (8.7.2)$$



For  $(i,j) = (1,2)$ , we have, by equation (8.3.5),

$$\begin{aligned} s_{12}(\mu) &= \sum_{\mu=0}^{\infty} \beta^{\mu} \sum_{m+n}^{\infty} \sum_{k=0}^{\mu-1} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v} \\ &= \frac{\beta n}{m+n} \sum_{\mu=0}^{\infty} \sum_{k=0}^{\infty} \beta^{\mu} \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v}. \end{aligned}$$

(8.7.3)

Since

$$\frac{n}{m+n} \left| \sum_{k=0}^{\infty} \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v} \right| = |E[(1-\bar{x}\bar{y})^k]| \leq 1, \quad (8.7.4)$$

we have

$$\begin{aligned} &\frac{\beta n}{m+n} \sum_{\mu=0}^{\infty} \sum_{k=0}^{\infty} \beta^{\mu} \left| \sum_{v=0}^k \binom{k}{v} (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v} \right| \\ &\leq \beta \sum_{\mu=0}^{\infty} \sum_{k=0}^{\infty} \beta^{\mu} = \beta \sum_{\mu=0}^{\infty} (\mu+1) \beta^{\mu}. \end{aligned} \quad (8.7.5)$$

The ratio test shows that  $\sum_{\mu=0}^{\infty} (\mu+1) \beta^{\mu}$  converges, hence, we may interchange the first two summation operators in (8.7.3) to obtain

$$s_{12}(\mu) = \frac{\beta n}{(1-\beta)(m+n)} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} \beta^k (-1)^v \frac{(n)_{k-v}(p)_v}{(m+n+1)_{k-v}(p+q)_v}. \quad (8.7.6)$$

Neglecting the constant multiplier, the series of absolute values corresponding to (8.7.6) is, upon interchanging the order of summation,

$$\begin{aligned} &\sum_{v=0}^{\infty} \sum_{k=0}^{\infty} \frac{(kv)! (n)_k (p)_v}{k! v! (m+n+1)_k (p+q)_v} \beta^k \beta^v \\ &= F_2(1, m, p, m+n+1, p+q; \beta, \beta), \end{aligned} \quad (8.7.7)$$

where  $F_2(a, b, \beta^1, \gamma, \gamma'; x, y)$  is Appell's second hypergeometric function of



two variables [2]. Appell has shown that the series (8.7.7) converges if  $0 \leq \beta < \frac{1}{2}$  and diverges if  $\frac{1}{2} < \beta < 1$ . The case  $\beta = \frac{1}{2}$  has not yet been investigated. We conclude that (8.7.6) converges absolutely for  $0 \leq \beta < \frac{1}{2}$  and, since Theorem 4.3.2 implies the convergence of (8.7.6), that the series converges conditionally for  $\frac{1}{2} < \beta < 1$ .

For  $(i,j) = (2,1)$  we use equation (8.3.7) to obtain

$$S_{21}^{(M)} = \frac{\beta}{1-\beta} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} \beta^k (-1)^v \frac{\binom{n}{k-v} \binom{p}{v+1}}{\binom{m+n}{k-v} \binom{p+q}{v+1}}, \quad (8.7.8)$$

the series converging absolutely for  $0 \leq \beta < \frac{1}{2}$  and conditionally for  $\frac{1}{2} < \beta < 1$ .

The remaining cases are

$$S_{11}^{(M)} = \sum_{\mu=1}^{\infty} \beta^{\mu} [1 - S_{12}^{(M)}] = \frac{\beta}{1-\beta} - S_{12}^{(M)} \quad (8.7.9)$$

and

$$S_{22}^{(M)} = \frac{\beta}{1-\beta} - S_{21}^{(M)}. \quad (8.7.10)$$

The expected discounted reward starting from state 1 is

$$\bar{V}_1^{(M)} = \sum_{k=1}^2 \bar{p}_{1k}^{(M)} r_{1k} + \sum_{j=1}^2 \sum_{k=1}^2 S_{1j}^{(M)} \bar{p}_{jk}^{(M)} r_{jk}. \quad (8.7.11)$$

Using the reward matrix (8.1.9) and the formulas (8.7.6) and (8.7.9), we obtain, upon collecting terms,

$$\begin{aligned} \bar{V}_1^{(M)} &= \frac{ma + nb}{(1-\beta)(m+n)} + \frac{n\beta}{(1-\beta)(m+n)} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} \beta^k (-1)^v \\ &\times \frac{\binom{n}{k-v} \binom{p}{v}}{\binom{m+n+1}{k-v} \binom{p+q}{v}} \left[ \frac{c(p+v) + dq}{p+q+v} - \frac{a(mk-v) + b(n+1)}{m+n+1+k-v} \right]. \end{aligned} \quad (8.7.12)$$

In a similar manner we find



$$\begin{aligned} \bar{V}_2(M) &= \frac{ep + dq}{(1-\beta)(p+q)} + \frac{\beta}{1-\beta} \sum_{k=0}^{\infty} \sum_{v=0}^k \binom{k}{v} \beta^k (-1)^v \frac{(m)_{k-v} (p)_{v+1}}{(m+n)_{k-v} (p+q)_{v+1}} \\ &\times \left[ \frac{a(m+k-v) + bn}{m+n+k-v} - \frac{c(p+v+1) + dq}{p+q+v+1} \right]. \end{aligned} \quad (8.7.13)$$

It can be shown that  $\bar{V}_1(M)$  and  $\bar{V}_2(M)$  satisfy equation (4.3.9).

### 8.8 A Generalization of a Result of Sher.

N. Z. Sher [37] has considered a game-theoretic model of a two-state Markov chain with alternatives and rewards. He shows that under certain circumstances each player should act so as to maximize his expected one-step transition reward. This result is generalized here.

Consider a two-state Markov chain with  $K_i$  alternatives in state  $i$  ( $i=1, 2$ ). Assume that the rewards depend only on the initial and final states  $i$  and  $j$  and not on the alternative used in making a transition from  $i$  to  $j$ . Assume further that the reward matrix is

$$\underline{R} = [r_{ij}^k] \quad (8.8.1)$$

where, if  $r$  is any real number,

$$r_{11}^k = r \quad k=1, \dots, K_1 \quad (8.8.2a)$$

$$r_{12}^k = r + \Delta_1 \quad k=1, \dots, K_1 \quad (8.8.2b)$$

$$r_{21}^k = r + \rho \quad k=1, \dots, K_2 \quad (8.8.2c)$$

$$r_{22}^k = r + \rho + \Delta_2 \quad k=1, \dots, K_2 \quad (8.8.2d)$$

We require that  $\rho \geq 0$  and  $\Delta_2 \geq \Delta_1 \geq 0$ .

Let  $\widetilde{P} = [\widetilde{p}_{ij}^k]$  be the matrix of alternative transition probabilities and let  $\widetilde{P}$  have the prior distribution function  $H(\widetilde{P}|\Gamma)$ . If  $F_{\Sigma}(P|\gamma)$



is the marginal distribution function of  $\tilde{P}(\underline{\sigma})$ , it is assumed that, for all  $\underline{\sigma} \in \Sigma$ ,  $F_{\underline{\sigma}}(P|\Gamma)$  is continuous on the boundary of  $\mathcal{S}_2$ .

The expected gain of the system under the policy  $\underline{\sigma}$  is  $\tilde{g}(\underline{\sigma}, \Gamma)$ . Suppose it is desired to choose a policy,  $\underline{\sigma}^*$ , which maximizes the expected gain,

$$\tilde{g}(\underline{\sigma}^*, \Gamma) = \max_{\underline{\sigma} \in \Sigma} \left\{ \tilde{g}(\underline{\sigma}, \Gamma) \right\}. \quad (8.8.3)$$

We shall show that, with the reward structure (8.8.2), it is sufficient to solve the corresponding deterministic problem for  $\bar{P}(\cdot | \Gamma) = E[\tilde{P}(\cdot | \Gamma)]$  and that the optimal policy,  $\underline{\sigma}^* = (\sigma_1^*, \sigma_2^*)$ , is determined by the equations

$$\tilde{p}_{12}^{(n)}(\Gamma) = \max_{k=1, \dots, K_1} \left\{ \tilde{p}_{12}^k(\Gamma) \right\}. \quad i=1, 2. \quad (8.8.4)$$

Let  $G_i^{(n)}(\underline{\sigma}, P)$  be the conditional total expected reward in  $n$  transitions under the policy  $\underline{\sigma}$  when the system starts from state  $i$  ( $i=1, 2$ ) and  $\tilde{P}(\underline{\sigma}) = P$ . Let  $\tilde{G}_i^{(n)}(\underline{\sigma}, \Gamma)$  be the corresponding unconditional expected reward,

$$\tilde{G}_i^{(n)}(\underline{\sigma}, \Gamma) = \int_{\mathcal{S}_2} G_i^{(n)}(\underline{\sigma}, P) dF_{\underline{\sigma}}(P | \Gamma). \quad (8.8.5)$$

$i=1, 2$   
 $n=1, 2, \dots$   
 $\underline{\sigma} \in \Sigma$

Lemma 8.8.1 For  $n=1, 2, \dots$  and all  $\underline{\sigma} \in \Sigma$ ,

$$\tilde{G}_1^{(n)}(\underline{\sigma}, \Gamma) - \tilde{G}_2^{(n)}(\underline{\sigma}, \Gamma) \leq \Delta_1. \quad (8.8.6)$$

Proof. We show that, for all  $P \in \mathcal{A}_2^*$ ,

$$G_1^{(n)}(\underline{\sigma}, P) - G_2^{(n)}(\underline{\sigma}, P) \leq \Delta_1, \quad n=1, 2, \dots \quad (8.8.7)$$

$\underline{\sigma} \in \Sigma$

from which equation (8.8.6) follows, since  $\mathcal{A}_2 - \mathcal{A}_2^*$  is a set of measure zero relative to  $F_{\underline{\sigma}}(P | \Gamma)$ .



Let  $\tilde{F}_{ij}(u, n)$  be the expected number of transitions from state  $i$  to state  $j$  in  $n$  transitions when the system starts from state  $u$ . Then, for all  $u \in S_2^*$ ,

$$G_i^{(n)}(\Sigma, P) = nr + p[\tilde{F}_{21}(1, n) + \tilde{F}_{22}(1, n)] + \Delta_1 \tilde{F}_{12}(1, n) + \Delta_2 \tilde{F}_{22}(1, n). \quad (8.8.8)$$

$i=1, 2$   
 $n=1, 2, \dots$   
 $\Sigma \in \Sigma$

If  $P$  is represented as in (8.1.1) and the eigenvalues of  $P$  are  $\lambda_1^{=1}$  and  $\lambda_2^{=1-x-y}$ , we can use the spectral representation of  $P$  given by (8.1.4) together with the expression for  $\tilde{F}_{ij}(u, n)$  given by equation (6.1.15) to obtain

$$G_1^{(n)}(\Sigma, P) - G_2^{(n)}(\Sigma, P) = \frac{1 - \lambda_2^n}{1 - \lambda_2} [-p + x\Delta_1 - (1-y)\Delta_2]. \quad (8.8.9)$$

Since  $p \geq 0$  and  $\Delta_2 \geq \Delta_1 \geq 0$ ,

$$\begin{aligned} G_1^{(n)}(\Sigma, P) - G_2^{(n)}(\Sigma, P) &\leq \frac{1 - \lambda_2^n}{1 - \lambda_2} (x+y-1)\Delta_1 \\ &= -\lambda_2 \frac{1 - \lambda_2^n}{1 - \lambda_2} \Delta_1. \end{aligned} \quad (8.8.10)$$

If  $0 \leq \lambda_2 < 1$ , then

$$-\lambda_2 \frac{1 - \lambda_2^n}{1 - \lambda_2} \Delta_1 \leq 0 \leq \Delta_1, \quad (8.8.11)$$

whence, if  $-1 < \lambda_2 < 0$ ,

$$\begin{aligned} -\lambda_2 \frac{1 - \lambda_2^n}{1 - \lambda_2} \Delta_1 &= -\lambda_2 \Delta_1 (1 + \lambda_2 + \lambda_2^2 + \dots + \lambda_2^{n-1}) \\ &\leq -\lambda_2 \Delta_1 < \Delta_1. \end{aligned} \quad (8.8.12)$$

In either case, we obtain (8.8.7). Q.E.D.



Lemma 8.8.2 For  $i=1,2$  and  $\underline{\sigma} \in \Sigma$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \tilde{G}_i^{(n)}(\underline{\sigma}, \dagger) = \tilde{g}(\underline{\sigma}, \dagger). \quad (8.8.13)$$

Proof. Since

$$G_i^{(n)}(\underline{\sigma}, \underline{p}) = \sum_{\alpha=1}^2 \sum_{\beta=1}^2 \tilde{x}_{\alpha\beta}^{(i,n)} r_{\alpha\beta}, \quad (8.8.14)$$

equation (6.1.14) yields

$$\tilde{G}_i^{(n)}(\underline{\sigma}, \dagger) = \sum_{\alpha=1}^2 \sum_{\beta=1}^2 r_{\alpha\beta} \sum_{k=0}^{n-1} E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] \quad (8.8.15)$$

where

$$E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] = \int_{\delta_a} p_{\alpha\beta} p_{1a}^{(k)} dF_{\underline{\sigma}}(p \mid \dagger). \quad (8.8.16)$$

Let  $\epsilon > 0$  be given. By a trivial extension of Theorem 4.2.5, there exists an integer  $v > 0$  such that, if  $k > v$ ,

$$\left| E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] - E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{\pi}_a] \right| < \frac{\epsilon}{2}. \quad (8.8.17)$$

Then, for  $n > v$ ,

$$\begin{aligned} & \left| E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{\pi}_a] - \frac{1}{n} \sum_{k=0}^{n-1} E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] \right| \\ & < \frac{1}{n} \sum_{k=0}^v \left| E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{\pi}_a] - E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] \right| + \frac{n-v}{n} \cdot \frac{\epsilon}{2} \\ & < \epsilon \end{aligned} \quad (8.8.18)$$

for  $n$  sufficiently large. Thus,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{p}_{1a}^{(k)}] = E_{\underline{\sigma}} [\tilde{p}_{\alpha\beta} \tilde{\pi}_a] \quad (8.8.19)$$

and, by (4.4.4),

$$\lim_{n \rightarrow \infty} \frac{1}{n} \tilde{G}_i^{(n)}(\underline{\sigma}, \dagger) = \tilde{g}(\underline{\sigma}, \dagger). \quad (8.8.20)$$

Q.E.D.

Theorem 8.8.3 Let  $\underline{\sigma}^* = (\sigma_1^*, \sigma_2^*)$  be a policy such that



$$\tilde{p}_{12}^{\sigma_1^0}(\gamma) = \max_{k=1, \dots, K_1} \left\{ \tilde{p}_{12}^{(k)}(\gamma) \right\}. \quad i=1, 2 \quad (8.8.21)$$

Then

$$\tilde{g}(\underline{\gamma}^*, \gamma) \geq \tilde{g}(\underline{\gamma}, \gamma). \quad \underline{\gamma} \in \Sigma \quad (8.8.22)$$

Proof. We first establish by induction that

$$\tilde{G}_1^{(n)}(\underline{\gamma}^*, \gamma) \geq \tilde{G}_1^{(n)}(\underline{\gamma}, \gamma). \quad \begin{matrix} i=1, 2 \\ n=1, 2, \dots \\ \underline{\gamma} \in \Sigma \end{matrix} \quad (8.8.23)$$

For  $n=1$ ,

$$\begin{aligned} \tilde{G}_1^{(1)}(\underline{\gamma}^*, \gamma) &= r + \tilde{p}_{12}^{\sigma_1^0}(\gamma) \Delta_1 \geq \tilde{G}_1^{(1)}(\underline{\gamma}, \gamma) \\ \tilde{G}_2^{(1)}(\underline{\gamma}^*, \gamma) &= r + p + \tilde{p}_{22}^{\sigma_2^0}(\gamma) \Delta_2 \geq \tilde{G}_2^{(1)}(\underline{\gamma}, \gamma). \quad (8.8.24) \\ &\quad \underline{\gamma} \in \Sigma \end{aligned}$$

Assume (8.8.23) holds for  $n$ . For  $i=1$ ,

$$\begin{aligned} \tilde{G}_1^{(n+1)}(\underline{\gamma}^*, \gamma) &= \tilde{p}_{11}^{\sigma_1^0}(\gamma) [r + \tilde{G}_1^{(n)}(\underline{\gamma}^*, \gamma)] + \tilde{p}_{12}^{\sigma_1^0}(\gamma) [r + \Delta_1 + \tilde{G}_2^{(n)}(\underline{\gamma}^*, \gamma)] \\ &= r + \tilde{p}_{12}^{\sigma_1^0}(\gamma) [\Delta_1 + \tilde{G}_2^{(n)}(\underline{\gamma}^*, \gamma) - \tilde{G}_1^{(n)}(\underline{\gamma}^*, \gamma)] + \tilde{G}_1^{(n)}(\underline{\gamma}^*, \gamma). \quad (8.8.25) \end{aligned}$$

Since  $\underline{\gamma}^*$  is an optimal policy for a transition interval of length  $n$ , we have, for all  $\underline{\gamma} \in \Sigma$ ,

$$\tilde{G}_1^{(n+1)}(\underline{\gamma}, \gamma) \leq r + \tilde{p}_{12}^{\sigma_1^0}(\gamma) [\Delta_1 + \tilde{G}_2^{(n)}(\underline{\gamma}, \gamma) - \tilde{G}_1^{(n)}(\underline{\gamma}, \gamma)] + \tilde{G}_1^{(n)}(\underline{\gamma}, \gamma) \quad (8.8.26)$$

and, by (8.8.21) and Lemma 8.8.1,

$$\begin{aligned} \tilde{G}_1^{(n+1)}(\underline{\gamma}, \gamma) &- \tilde{G}_1^{(n+1)}(\underline{\gamma}^*, \gamma) \\ &\geq [\tilde{p}_{12}^{\sigma_1^0}(\gamma) - \tilde{p}_{12}^{\sigma_1^0}(\gamma^*)] [\Delta_1 + \tilde{G}_2^{(n)}(\underline{\gamma}, \gamma) - \tilde{G}_1^{(n)}(\underline{\gamma}, \gamma)] \\ &\geq 0. \quad (8.8.27) \end{aligned}$$

Similarly,



$$\tilde{G}_2^{(n+1)}(\underline{\sigma}^*, +) = r + \rho + \frac{\sigma^*}{\tilde{p}_{22}^{(n)}} [\Delta_2 + \tilde{G}_2^{(n)}(\underline{\sigma}^*, +) - \tilde{G}_1^{(n)}(\underline{\sigma}^*, +)] + \tilde{G}_1^{(n)}(\underline{\sigma}^*, +), \quad (8.8.28)$$

and, since  $\Delta_2 \geq \Delta_1$ ,

$$\begin{aligned} \tilde{G}_2^{(n+1)}(\underline{\sigma}^*, +) &= \tilde{G}_2^{(n+1)}(\underline{\sigma}, +) \\ &\geq [\frac{\sigma^*}{\tilde{p}_{22}^{(n)}} - \tilde{p}_{22}^{(n)}] [\Delta_2 + \tilde{G}_2^{(n)}(\underline{\sigma}^*, +) - \tilde{G}_1^{(n)}(\underline{\sigma}^*, +)] \\ &\geq 0, \end{aligned} \quad \Sigma \in \Sigma \quad (8.8.29)$$

proving the induction.

Equation (8.8.23) and Lemma 8.8.2 together imply that, for all  $\underline{\sigma} \in \Sigma$ ,

$$\begin{aligned} \tilde{g}(\underline{\sigma}^*, +) &= \lim_{n \rightarrow \infty} \frac{1}{n} \tilde{G}_1^{(n)}(\underline{\sigma}^*, +) \\ &\geq \lim_{n \rightarrow \infty} \frac{1}{n} \tilde{G}_1^{(n)}(\underline{\sigma}, +) \\ &= \tilde{g}(\underline{\sigma}, +). \end{aligned} \quad (8.8.30)$$

Q.E.D.



## CHAPTER 9

### CONCLUDING REMARKS

In the foregoing chapters we have described a formal structure for certain broad classes of sequential sampling and fixed sample-size decision problems in a Markov chain with unknown transition probabilities. Since there is very little theory in this area most of our efforts have been directed toward answering questions of existence and convergence. For this reason the portions of this report which deal with numerical computation set forth the obvious, but not necessarily the most efficient, ways to approach problems of calculation. It does seem clear, however, that, for problems with a large number of states in which a high degree of accuracy is required, we must think in terms of hours, not minutes, of computer time. This is not to say that the Bayesian method of dealing with Markov chains with uncertain transition probabilities must be abandoned as impractical. But it must be recognised that, at the present state of the art, the Bayesian treatment is probably most practical for problems with two or three states, loose prior distributions, and large differences in the rewards associated with different actions. As problems tend to differ from these criteria, the decision-maker must balance increasing computation time against the required accuracy of the solution and choose an appropriate approximation.

There are numerous questions of immediate interest which remain unanswered, some theoretical and some numerical. Many of these are listed below.



1. Certain error bounds were derived in Chapters 3, 4, and 5 which depend on the discount factor,  $\beta$ , but not on  $\gamma$ , the parameter of the prior distribution. These bounds should be made tighter for specific prior distributions by including factors which involve  $\gamma$ .
2. The rate of convergence of the successive approximations methods developed in Chapters 3, 4, and 5 depend upon the choice of terminal functions. Classes of terminal functions which accelerate this convergence rate should be investigated.
3. The analysis of undiscounted adaptive control models by letting  $\beta \rightarrow 1$  in the corresponding discounted problem may provide a workable approach to a difficult matter. The remarks of Section 3.6 are relevant in this connection.
4. The question of the uniqueness of solutions to equations (4.2.40) and (4.2.55) is of considerable importance for the calculation of the means and product moments of the steady-state vector  $\tilde{\pi}$  when a method of successive approximations is used. The problem of the convergence of the approximant  $\tilde{\pi}(n, \gamma)$ , as defined by equation (4.2.42) with the terminal function (4.2.51), is also of importance.
5. In the terminal control models of Chapter 5 it is necessary to evaluate expressions of the form

$$\tilde{v}_1(\underline{\sigma}^*, \gamma) = \max_{\underline{\sigma} \in \Sigma} \left\{ \tilde{v}_1(\underline{\sigma}, \gamma) \right\}$$

and

$$\tilde{g}(\underline{\sigma}^*, \gamma) = \max_{\underline{\sigma} \in \Sigma} \left\{ \tilde{g}(\underline{\sigma}, \gamma) \right\} .$$



At present the only method of finding the maximizing policy  $\underline{\sigma}^*$  is by direct search over the elements of  $\Sigma$ . More efficient methods of finding  $\underline{\sigma}^*$  should be investigated; approximations to  $\underline{\sigma}^*$  of the sort described in Section 5.4 should also be studied.

6. A formal analysis of the undiscounted terminal control models III and IV, which were introduced in Section 5.5, should be carried out. This analysis would examine such questions as the existence and uniqueness of solutions, the convergence of successive approximation methods, and whether a terminal decision point is reached with probability one in an optimal sampling strategy. In this regard it is to be noted that equations (5.5.2) and (5.5.3) can be made more precise by replacing the expression

$$\max_{\underline{\sigma} \in \Sigma} \{ \bar{v}_g(\underline{\sigma}, \tau) \}$$

by the expression

$$\max_{\underline{\sigma} \in \Sigma} \left\{ \bar{w}_1(\underline{\sigma}, \tau) + \bar{v}_g(\underline{\sigma}, \tau) \right\},$$

where  $\bar{w}_1(\underline{\sigma}, \tau)$  is the expected relative value<sup>\*</sup> of starting the system in state 1 and operating it indefinitely under the policy  $\underline{\sigma}$  when the prior distribution function of  $\underline{\Phi}$  is  $H(\underline{\Phi} | \tau)$ . Methods of computing  $\bar{w}_1(\underline{\sigma}, \tau)$  have not yet been studied.

7. There are well-known difficulties in assigning a multivariate prior distribution to the elements of  $\underline{\Phi}$  in such a manner as to accurately reflect the decision-maker's state of knowledge. It would be of considerable interest, therefore, to investigate the sensitivity of some of the models of foregoing chapters to relatively small changes in the prior distribution.

---

\* Cf. Howard [22], Ch. 4.



In addition to these and other immediate questions which arise in connection with the research reported in this study, there are several fairly obvious directions in which this research can be extended. For example, many of the results and techniques developed here should be capable of extension to decision problems in a semi-Markov chain in which both the transition probabilities and the parameters of the holding-time distributions are uncertain.

More general stochastic processes should be amenable to Bayesian analysis, although different techniques than those utilised here may be required. The Weiner process, for example, can be analyzed with the existing Bayesian theory for normal processes.



APPENDIX A  
GLOSSARY OF SYMBOLS

<u>Symbol</u>	<u>Meaning</u>	<u>Defined on Page</u>
$B(p,q)$	Beta function.	---
$B_N(\alpha)$	Generalized beta function.	170
$\beta$	Discount factor.	38
$c_i$	Sampling cost when system is observed in state $i$ .	118
$\bar{c}_i^k(\tau)$	Expected one-step sampling cost when process is in state $i$ and alternative $k$ is to be used.	119
$C$	Maximum sampling cost.	124
$\text{cov}[\cdot]$ or $\text{cov}[\cdot \cdot]$	Covariance operator.	---
$e_j(n, \tau)$	Error of $n$ th successive approximation in adaptive and terminal control problems.	52 127 135
$e_j(\Sigma, \tau)$	Error of the terminal decision $\Sigma$ .	142 144
$E[\cdot]$ or $E[\cdot \cdot]$	Expectation operator.	---
$f_{\cdot i}$	Sum of $i$ th row of transition count.	151
$f_{i \cdot}$	Sum of $i$ th column of transition count.	151
$f_{ij}(u,n)$	Number of transitions from state $i$ to state $j$ in $n$ transitions when process starts from state $u$ .	155
$\tilde{f}_{ij}(u,n)$	Expected value of $\tilde{f}_{ij}(u,n)$ .	155
$f_B^{(N)}(p m)$	Multivariate beta probability density function.	170
$f_{p^*}^{(N)}(p   e, g)$	Nonstandardized multivariate beta probability density function.	176



Symbol	Meaning	Defined on Page
$f_{\text{BW}}^{(N)}(\underline{F} \mid \underline{u}, \underline{n}, \underline{M})$	Beta-Whittle probability mass function.	192
$f_{\text{BW2}}^{(N)}(\underline{F} \mid \underline{p}, \underline{n}, \underline{M})$	Beta-Whittle-2 probability mass function.	195
$f_{\text{BW2e}}^{(N)}(\underline{F} \mid \underline{n}, \underline{M})$	Nonstandard beta-Whittle-2 mass function.	196
$f_{\text{MB}}^{(K, N)}(\underline{P} \mid \underline{M})$	Matrix beta probability density function.	180
$f_{\text{MBe}}^{(K, N)}(\underline{P} \mid \underline{A}, \underline{M})$	Nonstandardized matrix beta probability density function.	186
$f_{\text{MB1}}^{(N)}(\underline{P} \mid \underline{M}, \underline{\lambda})$	Matrix beta-1 probability density function.	190
$f_{\text{W}}^{(N)}(\underline{F} \mid \underline{u}, \underline{n}, \underline{P})$	Whittle probability mass function	153
$f_{\text{W1}}^{(N)}(\underline{u}, \underline{F} \mid \underline{p}, \underline{n}, \underline{P})$	Whittle-1 probability mass function.	161
$f_{\text{W2}}^{(N)}(\underline{F} \mid \underline{p}, \underline{n}, \underline{P})$	Whittle-2 probability mass function.	162
$\underline{F} = [f_{ij}]$	Transition count.	151
$F_{\beta}^{(N)}(\underline{p} \mid \underline{m})$	Multivariate beta probability distribution function.	174
$F_{\Sigma}(\underline{P} \mid \dot{\gamma})$ or $F(\underline{P} \mid \dot{\gamma})$	Marginal distribution function of the $N$ rows of $\underline{F}$ specified by $\Sigma$ .	8
$F_2(a, \beta, \beta^*, \gamma, \gamma^*; x, y)$	Appell's second hypergeometric function of two arguments.	---
$\mathcal{J}$	Family of probability distribution functions, $F(\underline{P} \mid \dot{\gamma})$ .	34
$s(P)$	Expected reward per transition in the steady state, or process gain, when $\underline{P} = P$ .	108
$\tilde{g}(\dot{\gamma})$ or $\tilde{g}(\Sigma, \dot{\gamma})$	Expected gain under the policy $\Sigma$ .	109
$\Gamma(x)$	Gamma function.	---
$H(\underline{P} \mid \dot{\gamma})$	Probability distribution function for the generalized stochastic matrix $\underline{P}$ .	8



Symbol	Meaning	Defined on Page
$\mathcal{H}$	Family of probability distribution functions, $H(\underline{\theta} \Psi)$ .	11
$L_{\mathcal{B}_n}(\underline{\theta})$	Likelihood function for the sample $\underline{x}_n$ .	11
$\underline{P} = [p_{ij}^k]$	A $K \times N$ generalized stochastic matrix ( $K \geq N$ ).	7
$\underline{P}(\Sigma) = \underline{P} = [p_{ij}]$	An $N \times N$ stochastic matrix consisting of $N$ rows of $\underline{P}$ specified by $\Sigma$ .	7
$\underline{P}^T$	Transpose of the matrix $\underline{P}$ .	---
$\bar{p}_{ij}^k(\Psi)$	Expected value of $\hat{p}_{ij}^k$ .	41
$\bar{p}_{ij}^{(n)}$	$n$ -step transition probability when $\underline{P} = \underline{P}$ .	69
$\bar{p}_{ij}^{(n)}(\Psi) = \bar{p}_{ij}^{(n)}(\Sigma, \Psi)$	Expected $n$ -step transition probability under the policy $\Psi$ .	69
$\bar{\pi} = (\bar{\pi}_1, \dots, \bar{\pi}_N)$	Steady-state probability vector of an ergodic Markov chain when $\underline{P} = \underline{P}$ .	73
$\bar{\pi}(\Psi) = (\bar{\pi}_1(\Psi), \dots, \bar{\pi}_N(\Psi))$	Expected steady-state probability vector.	73
$\bar{\pi}(n, \Psi) =$		
$(\bar{\pi}_1(n, \Psi), \dots, \bar{\pi}_N(n, \Psi))$	The $n$ th successive approximation to $\bar{\pi}(\Psi)$ .	87
$\bar{\pi}_{ij}(\Psi)$	Expected value of $\hat{\pi}_i \hat{\pi}_j$ .	94
$\Phi_n(u, v, n, \underline{P})$	Set of all transition counts of size $n$ which start in state $u$ and end in state $v$ when $\underline{P}$ is the matrix of transition probabilities.	152
$\Phi_n(u, n, \underline{P})$	Set of all transition counts of size $n$ which start in state $u$ when $\underline{P}$ is the matrix of transition probabilities.	152
$\Phi_n^*(n, \underline{P})$	Set of all transition counts of size $n$ when $\underline{P}$ is the matrix of transition probabilities.	162



Syntol	Meaning	Defined on Page
$\Phi_{N1}^*(n, P)$	Set of all transition counts of size $n$ which start and end in the same state when $P$ is the matrix of transition probabilities.	163
$\Phi_{N2}^*(n, P)$	Set of all transition counts of size $n$ which start and end in different states when $P$ is the matrix of transition probabilities.	163
$\Phi_N(u, v, n)$	Set of all transition counts of size $n$ which start in state $u$ and end in state $v$ when the matrix of transition probabilities is positive.	192
$\Phi_N(u, n)$	Set of all transition counts of size $n$ which start in state $u$ when the matrix of transition probabilities is positive.	192
$\Phi_N^*(n)$	Set of all transition counts of size $n$ when the matrix of transition probabilities is positive.	195
$\Phi_{N1}^*(n)$	Set of all transition counts of size $n$ which start and end in the same state when the matrix of transition probabilities is positive.	195
$\Phi_{N2}^*(n)$	Set of all transition counts of size $n$ which start and end in different states when the matrix of transition probabilities is positive.	195
$\gamma$	Generic symbol for the parameters of a probability distribution function.	8
$\Psi$	Admissible parameter set.	11
$\Omega = [q_{ij}^k]$	True state of nature.	23
$\bar{q}_i^k(\gamma)$	Expected one-step transition reward when the system is in state $i$ and alternative $k$ is to be used.	41
$q_{ij}^{(n)}(\Sigma)$	The $n$ -step transition probability under the policy $\Sigma$ when $\Omega$ is the true state of nature.	24



Symbol	Meaning	Defined on Page
$\bar{q}_i^{(n)}(\Sigma, \beta, t)$ or $\bar{q}_i^{(n)}(\beta, t)$	Expected discounted reward in n transitions when the system starts in state i.	71
$\underline{R} = [r_{ij}^k]$	The K x N matrix of one-step transition rewards.	7
$\underline{R}(\Sigma) = R = [r_{ij}]$	The N x N matrix of one-step transition rewards consisting of the N rows of $\underline{R}$ specified by $\Sigma$ .	7
R	Maximum element of $\underline{R}$ .	51
r	Minimum element of $\underline{R}$ .	51
$\underline{R}^*$	The element of $\underline{R}$ with the largest absolute value.	45
$\underline{r}^*$	The element of $\underline{R}$ with the smallest absolute value.	45
$R_N(a)$	Range set of a random vector with the nonstandardized multivariate beta distribution.	176
$R_{K,N}(a)$	Range set of a random matrix with the nonstandardised matrix beta distribution.	185
$\mathcal{S}_{K,N}$	Set of all K x N generalized stochastic matrices.	8
$\mathcal{S}_N$	Set of all N x N stochastic matrices.	8
$\mathcal{S}_N^+$	Set of all N x N positive stochastic matrices.	74
$\mathcal{S}_N^a$	Set of all N x N stochastic matrices with elements in the closed interval [a, 1-a].	74
$\Sigma = (\tau_1, \dots, \tau_N)$	Policy vector.	7
$\tau_{\alpha\beta\gamma\delta}(u,n)$	Expected value of $\tilde{f}_{\alpha\beta}(u,n)\tilde{f}_{\gamma\delta}(u,n)$	158
$\Sigma$	Set of all policy vectors, $\Sigma$ .	8



Symbol	Meaning	Defined on Page
$\tau_{ij}^k(\gamma)$	Parameter of the posterior distribution of $\underline{P}$ when the parameter of the prior distribution is $\gamma$ and a transition from state i to state j under alternative k is observed.	12
$\tau_{ij}(\gamma)$	Parameter of the posterior distribution of $\underline{P}$ when the parameter of the prior distribution is $\gamma$ and a transition from state i to state j is observed.	12
$v_{ij}(n, \underline{\sigma}, \gamma)$	Parameter of the posterior distribution of $\underline{P}$ when the parameter of the prior distribution is $\gamma$ , the system starts in state i and is observed in state j after n transitions under the policy $\underline{\sigma}$ .	130
$v_i(\gamma)$	Expected total discounted reward over an infinite period when the system starts in state i and an optimal sampling strategy is followed.	40 118 130
$v_i(n, \gamma)$	The nth successive approximation to $v_i(\gamma)$	46 129 134
$v_i(\gamma; v)$	Expected total reward over a period with terminal operation phase of length v when the system starts from state i and an optimal sampling strategy is followed.	145 146
$v_i(\gamma, \underline{\sigma})$	Expected total discounted reward over an infinite period when the system starts in state i with the policy $\underline{\sigma}$ in force and an optimal sampling strategy is followed (discounted process with set-up cost).	148
$v^*$	Minimum of a set of constant terminal reward functions.	51
$v^*$	Maximum of a set of constant terminal reward functions.	51
$v_i(\gamma)$	Terminal reward function.	46
$V$	Bound on the terminal reward functions.	47



Symbol	Meaning	Defined on Page
$V_1(P)$	Expected total discounted reward over an infinite period when $\tilde{P} = P$ , the system starts from state 1, and a fixed policy is used.	96
$\tilde{V}_1(\underline{\pi}, t)$ or $\tilde{V}_1(t)$	Expected total discounted reward over an infinite period when the system starts from state 1 and the policy $\underline{\pi}$ is used.	97
$\tilde{V}_1(n, t)$	The $n$ th successive approximation to $V_1(t)$ .	100
$\text{var}[\cdot]$ or $\text{var}[\cdot \quad]$	The variance operator.	---
$x_n = (x_0, x_1, \dots, x_n)$	A sample of $n + 1$ states occupied by a Markov chain.	11



## APPENDIX B

### PROGRAM VITERATION TO SOLVE EQUATIONS (3.2.1) AND (3.2.2).

PROGRAM NAME IS VITERATION.  
THIS PROGRAM RECURSIVELY COMPUTES VALUES OF V(I, T<sub>i</sub>, M) FOR  
R<sub>i</sub>=1,...,N AND T<sub>i</sub>=S,...,T, FOLLOWING AN OPTIMAL POLICY. THE  
REWARD MATRIX IS R AND THE TERMINAL REWARD VECTOR IS RHO.  
THE MAXIMIZATION IS OVER THE MU(I) ALTERNATIVES IN STATE I.  
BETA IS THE DISCOUNT FACTOR. A MATRIX BETA PRIOR IS  
ASSUMED.

```
PROGRAM COMMON R,RHO,RDIM,IND,MU,N,BETA,IND1,LIST,TOP,  
OINDIM,POL  
  INTEGER IND,MU,N,IND1,TOP,S,T,I,J,K,MAXSP,N1,POL,V2  
  DIMENSION R(500,RDIM),M(500,RDIM),RHO(10),IND(100,INDIM),  
  OMU(10),IND1(10),LIST(21000),V1(10),V2(10)  
  VECTOR VALUES RDIM=3,1,0,0  
  VECTOR VALUES INDIM=2,1,0  
READ READ FORMAT INPUT1, N,S,T,MAXSP,BETA  
PRINT FORMAT OUT1A, BETA  
RDIM(2)=N  
RDIM(3)=N  
INDIM(2)=N  
IND1(1)=0  
IND(1,1)=0  
THROUGH ALFA1, FOR K=1,1,K.E.N  
ALFA1 IND(1,K+1)=K*N  
N1=N*N  
THROUGH ALFA, FOR K=1,1,K.E.MAXSP  
IND1(K+1)=K*N  
IND(K+1,1)=K*N1  
THROUGH ALFA, FOR I=1,1,I.E.N  
ALFA IND(K+1,I+1)=K*N1+I*N  
READ FORMAT INPUT2, MU(1) ... MU(N)  
PRINT FORMAT OUT1E, (I=1,1,I.G.N, MU(I))  
READ FORMAT INPUT3, R(1,1,1) ... R(MU(N),N,N), M(1,1,1)  
0... M(MU(N),N,N), RHO(1) ... RHO(N)  
PRINT FORMAT OUT1D, (I=1,1,I.G.N, RHO(I))  
PRINT FORMAT OUT1B, (I=1,1,I.G.N, (K=1,1,K.G.MU(I),  
0(J=1,1,J.G.N, M(IND(IND1(K)+I)+J))))  
PRINT FORMAT OUT1C, (I=1,1,I.G.N, (K=1,1,K.G.MU(I),  
0(J=1,1,J.G.N, R(IND(IND1(K)+I)+J))))  
SET LIST TO LIST  
LIST = 0  
THROUGH DELTA, FOR K=S,1,K.G.T  
THROUGH GAMMA, FOR I=1,1,I.G.N  
V1(I)=VMAX.(I,K,M)  
GAMMA V2(I)=POL
```



```
PRINT FORMAT OUT2, K, V1(1) ... V1(N)
DELTA PRINT FORMAT OUT3, V2(1) ... V2(N)
TRANSFER TO READ
```

RFORMAT SPECIFICATIONS

```
VECTOR VALUES INPUT1=$4I10,F10.5*$
VECTOR VALUES INPUT2=$10I7*$
VECTOR VALUES INPUT3=$(7F10.5)*$*
VECTOR VALUES OUT1A=$7H1BETA =,G15.5*$*
VECTOR VALUES OUT1B=$6HO M =,8G15.5/(8G15.5)*$*
VECTOR VALUES OUT1C=$6HO R =,8G15.5/(8G15.5)*$*
VECTOR VALUES OUT1D=$6HORHO =,8G15.5/8G15.5*$*
VECTOR VALUES OUT1E=$5HOMU =,10I5*$*
VECTOR VALUES OUT2=$8H0FOR T =,I2,I4H V(I, T, M) =,7G15.5/
08G15.5*$*
VECTOR VALUES OUT3=$7H POLICY,10I5*$*
END OF PROGRAM
```

EXTERNAL FUNCTION (I1, N1, M)  
ENTRY TO VMAX.

RTHIS FUNCTION RECURSIVELY COMPUTES MAX V(I1,N1,M)=Y, THE  
RMAXIMUM EXPECTED RETURN IN N1 STEPS IF THE SYSTEM STARTS IN  
RSTATE I1 WITH PARAMETER MATRIX M. PRIOR DISTRIBUTION IS  
RMATRIX BETA. MAXIMIZATION IS OVER THE MU(I1)  
RALTERNATIVES IN STATE I1.

```
PROGRAM COMMON R,RHO,RDIM,IND,MU,N,BETA,IND1,LIST,TOP,
OINDIM,POL
INTEGER I1,N1,I,N2,K,IND,MU,N,RDIM,J,IND1,LIST,POl
DIMENSION R(500,RDIM),RHO(10),RDIM(3),IND(100,INDIM),
O LIST(21000),MU(10),IND1(10),INDIM(2),TM(500,RDIM)
I=I1
N2=N1
Y=1.E-35
WHENEVER N2.E.0, FUNCTION RETURN RHO(I)
THROUGH ALFA, FOR K=1,1,K.G.MU(I)
MSUM=0.
THROUGH PHI, FOR J=1,1,J.G.N
PHI MSUM=MSUM+M(IND(IND1(K)+I)+J)
STOR=0.
THROUGH GAMMA, FOR J=1,1,J.G.N
SAVE RETURN
SAVE DATA N2,POl,MSUM,STOR,Y,M(K,I,J) ... M(MU(I),I,N),I,J,K
EXECUTE TR1.(I,J,K,M,TM)
X=VMAX.(J,N2-1,TM)
```



```
RESTORE DATA K,J,I,M(MU(I),I,N) ... M(K,I,J),Y,STOR,MSUM,
OPOL,N2
RESTORE RETURN
GAMMA STOR=STOR+(M(IND(IND1(K)+I)+J)/MSUM)*(R(IND(IND1(K)+I)+J)
0+BETA*X)
WHENEVER STOR .LE. Y, TRANSFER TO ALFA
Y=STOR
POL=K
ALFA CONTINUE
FUNCTION RETURN Y
END OF FUNCTION
```

EXTERNAL FUNCTION (I1,J1,K1,M,TM)  
ENTRY TO TR1.

RTHIS FUNCTION EFFECTS THE TRANSFORMATION FROM THE PRIOR  
RPARAMETER MATRIX M TO THE POSTERIOR PARAMETER MATRIX  
RTR1.(I1,J1,K1,M)=TM, WHEN A TRANSITION IS OBSERVED FROM I1 TO  
RJ1 UNDER ALTERNATIVE K1. PRIOR DISTRIBUTION IS MATRIX BETA.

```
PROGRAM COMMON R,RHO,RDIM,IND,MU,N,BETA,IND1,LIST,TOP,
OINDIM,POL
INTEGER I1,J1,K1,I,J,K,IND,MU,N,IND1
DIMENSION R(500,RDIM),RHO(10),RDIM(3),IND(100,INDIM),
OLIST(21000),MU(10),IND1(10),INDIM(2)
THROUGH ALFA, FOR I=1,1,I.G.N
THROUGH ALFA, FOR J=1,1,J.G.N
THROUGH ALFA, FOR K=1,1,K.G.MU(I)
ALFA TM(IND(IND1(K)+I)+J)=M(IND(IND1(K)+I)+J)
TM(IND(IND1(K1)+I1)+J1)=M(IND(IND1(K1)+I1)+J1)+1.0
FUNCTION RETURN
END OF FUNCTION
```



## APPENDIX C

### PROGRAM PHI MATRIX TO COMPUTE EQUATION (4.1.2)

RPROGRAM NAME IS PHI MATRIX  
RTHIS PROGRAM RECURSIVELY COMPUTES VALUES OF PHI(I,J,T1,M)  
RFOR I,J=1,...,N AND T1=S,...,T. A MATRIX BETA  
RPRIOR IS ASSUMED.

PROGRAM COMMON IND,N,J,LIST,TOP,MDIM  
INTEGER N,IND,I,J,S,T,K,LIST  
DIMENSION M(100,MDIM),IND(10),LIST(21000)\*F(10)  
VECTOR VALUES MDIM=2,1,0  
READ READ FORMAT INPUT1, N,S,T  
MDIM(2)=N  
IND(1)=0  
THROUGH ALFA, FOR K=1,1,K.E.N  
ALFA IND(K+1)=K\*N  
READ FORMAT INPUT2, M(1,1) ... M(N,N)  
PRINT FORMAT OUT1, N,S,T, (K=1,1,K.G.N,(L=1,1,L.G.N,  
OM(IND(K)+L)))  
SET LIST TO LIST  
LIST=0  
THROUGH GAMMA, FOR K=S,1,K.G.T  
THROUGH GAMMA, FOR I=1,1,I.G.N  
THROUGH DELTA, FOR J=1,1,J.E.N  
DELTA F(J)=PHI.(I,K,M)  
F(N)=1.0  
THROUGH EPS, FOR J=1,1,J.E.N  
EPS F(N)=F(N)-F(J)  
WHENEVER I.E.1  
PRINT FORMAT OUT2, K, F(1) ... F(N)  
OTHERWISE  
PRINT FORMAT OUT3, F(1) ... F(N)  
GAMMA END OF CONDITIONAL  
TRANSFER TO READ

RFORMAT SPECIFICATIONS  
VECTOR VALUES INPUT1=\$3I10\*\$  
VECTOR VALUES INPUT2=\$(7F10.5)\*\$  
VECTOR VALUES OUT1=\$3H1N=,I5,4H S=,I5,4H T=,I5/  
0(1H , 8G15.5)\*\$  
VECTOR VALUES OUT2=\$7H0FOR T=, I2, 15H PHI(I,J,T,M)=,  
06G15.5\*\$  
VECTOR VALUES OUT3=\$S24, 6G15.5\*\$  
END OF PROGRAM



EXTERNAL FUNCTION (I1, T1, M)  
ENTRY TO PHI.

RTHIS FUNCTION RECURSIVELY COMPUTES PHI(I1,J,T1,M)=Y, THE  
RPROBABILITY THAT AT TIME T1 THE SYSTEM WILL BE IN STATE J,  
RGIVEN THAT AT TIME 0 IT WAS IN STATE I1 WITH PARAMETER  
RMATRIX M. PRIOR IS MATRIX BETA.

```
PROGRAM COMMON IND,N,J,LIST,TOP,MDIM
INTEGER I1,J,T1,I,T,K,N,IND,LIST,MDIM
DIMENSION IND(10),LIST(21000),MDIM(2),TM(100,MDIM)
I=I1
T=T1
MSUM=0.
THROUGH ALFA, FOR K=1,1,K.G.N
ALFA MSUM=MSUM+M(IND(I)+K)
WHENEVER T.E.1, FUNCTION RETURN M(IND(I)+J)/MSUM
Y=0.
THROUGH BETA, FOR K=1,1,K.G.N
SAVE RETURN
SAVE DATA Y,T,MSUM,M(1,1) ... M(N,N),I,K
EXECUTE TR.(I,K,M,TM)
X=PHI.(K,T-1,TM)
RESTORE DATA K,I,M(N,N) ... M(1,1),MSUM,T,Y
RESTORE RETURN
BETA Y=Y+(M(IND(I)+K)/MSUM)*X
FUNCTION RETURN Y
END OF FUNCTION
```

EXTERNAL FUNCTION (I,K,M,TM)
ENTRY TO TR.

RTHIS FUNCTION EFFECTS THE TRANSFORMATION FROM THE PRIOR  
RPARAMETER MATRIX M TO THE POSTERIOR PARAMETER MATRIX  
RT.(I,K,M)=TM, WHEN ONE TRANSITION FROM I TO K IS OBSERVED.  
RPRIOR IS MATRIX BETA.

```
PROGRAM COMMON IND,N,J,LIST,TOP,MDIM
INTEGER I,K,IND,J,L,N,J1,MDIM,LIST,TOP
DIMENSION IND(10),MDIM(2),LIST(21000)
THROUGH ALFA, FOR J1=1,1,J1.G.N
THROUGH ALFA, FOR L=1,1,L.G.N
ALFA TM(IND(J1)+L)=M(IND(J1)+L)
TM(IND(I)+K)=TM(IND(I)+K)+1.0
FUNCTION RETURN
END OF FUNCTION
```



## APPENDIX D

### PROGRAM PIAPROX TO COMPUTE EQUATIONS (4.2.42).

RPROGRAM NAME IS PIAPROX. THIS PROGRAM RECURSIVELY COMPUTES RVALUES OF THE SUCCESSIVE APPROXIMANT PI(I,T1,M) FOR RI=1,...,N AND T1=S,...,T. A MATRIX BETA PRIOR IS USED.

```
PROGRAM COMMON IND,N,LIST,MDIM,N1,ADIM,AIND
INTEGER N,IND,I,K,S,T,N1,AIND
DIMENSION M(100,MDIM),IND(10),LIST(21000),F(10),AIND(10)
VECTOR VALUES MDIM=2,1,0
VECTOR VALUES ADIM=2,1,0
READ READ FORMAT INPUT1, N,S,T
MDIM(2)=N
N1=N+1
ADIM(2)=N1
AIND(1)=0
IND(1)=0
THROUGH ALFA, FOR K=1,1,K.E.N
AIND(K+1)=K*N1
ALFA IND(K+1)=K*N
READ FORMAT INPUT2, M(1,1) ... M(N,N)
PRINT FORMAT OUT1, N,S,T,(K=1,1,K.G.N, (I=1,1,I.G.N,
OM(IND(K)+I)))
SET LIST TO LIST
LIST=0
THROUGH GAMMA, FOR K=S,1,K.G.T
THROUGH DELTA, FOR I=1,1,I.G.N
DELTA F(I)=PI.(I,K,M)
PRINT FORMAT OUT2, K, F(1) ... F(N)
SUM=0.
THROUGH BETA, FOR I=1,1,I.G.N
BETA SUM=SUM+F(I)
THROUGH EPS, FOR I=1,1,I.G.N
EPS F(I)=F(I)/SUM
PRINT FORMAT OUT3, F(1) ... F(N)
GAMMA PRINT FORMAT OUT4, SUM
TRANSFER TO READ
```

#### RFORMAT SPECIFICATIONS.

```
VECTOR VALUES INPUT1=$3I10*$
VECTOR VALUES INPUT2=$(7F10.5)*$
VECTOR VALUES OUT1=$3H1N=,I5,4H S=,I5,4H T=,I5/
0(1H ,8G15.5)*$
VECTOR VALUES OUT2=$7H0FOR T=,I2,10H PI(T,M)=,(6G15.5)*$
VECTOR VALUES OUT3=$19H NORMALIZED VECTOR=,(6G15.5)*$
VECTOR VALUES OUT4=$1H ,S11,7HC(T,M)=,G15.6*$
END OF PROGRAM
```



EXTERNAL FUNCTION(J1,T1,M)  
ENTRY TO PI.

RTHIS FUNCTION RECURSIVELY COMPUTES PI(J1,T1,M), THE T1TH  
RSUCCESSIVE APPROXIMANT TO THE J1TH ELEMENT OF THE MEAN  
RSTEADY-STATE PROBABILITY VECTOR WHEN THE PRIOR IS MATRIX  
RBETA WITH PARAMETER M.

```
PROGRAM COMMON IND,N,LIST,MDIM,N1,ADIM,AIND
INTEGER J1,T1,I,J,K,N,T,IND,MDIM,N1,ADIM,AIND
DIMENSION IND(10),LIST(21000),MDIM(2),TM(100,MDIM),PBAR(10),
          ADIM(2)*AIND(10)
J=J1
T=T1
THROUGH ALFA, FOR K=1,1,K.G.N
MSUM=0.
THROUGH BETA, FOR I=1,1,I.G.N
BETA MSUM=MSUM+M(IND(K)+I)
ALFA PBAR(K)=M(IND(K)+J)/MSUM
Y=0.
THROUGH GAMMA, FOR K=1,1,K.G.N
SAVE RETURN
SAVE DATA Y,T,PBAR(K) ... PBAR(N),K,J, M(1,1) ... M(N,N)
M(IND(K)+J)=M(IND(K)+J)+1.
WHENEVER T.G.1, TRANSFER TO ZETA
X=PIZERO.(K,M)
TRANSFER TO ETA
ZETA X=PI.(K,T=1,M)
ETA RESTORE DATA M(N,N) ... M(1,1),J,K,PBAR(N) ... PBAR(K),T,Y
RESTORE RETURN
GAMMA Y=Y+X*PBAR(K)
FUNCTION RETURN Y
END OF FUNCTION
```

EXTERNAL FUNCTION(L1,M)  
ENTRY TO PIZRO.

RTHIS FUNCTION COMPUTES THE TERMINAL FUNCTION PI(L1,0,M)  
RAS THE L1TH ELEMENT OF THE STEADY-STATE PROBABILITY VECTOR  
RCORRESPONDING TO THE MEAN OF THE PRIOR DISTRIBUTION.  
RPRIOR IS MATRIX BETA WITH PARAMETER M.

```
PROGRAM COMMON IND,N,LIST,MDIM,N1,ADIM,AIND
INTEGER L1,L,N,I,J,K,N1,IND,MDIM,ADIM,AIND
DIMENSION IND(10),LIST(21000),MDIM(2)*ADIM(2)*A(110,ADIM),
```



```
OAIND(10)
L=L1
THROUGH ALFA, FOR I=1,1,I.G.N
MSUM=0.
THROUGH BETA, FOR K=1,1,K.G.N
BETA MSUM=MSUM+M(IND(I)+K)
THROUGH GAMMA, FOR K=1,1,K.E.N
GAMMA A(AIND(K)+I)=-M(IND(I)+K)/MSUM
A(AIND(N)+I)=1.
ALFA A(AIND(I)+N1)=0.
A(AIND(N)+N1)=1.
THROUGH DELTA, FOR K=1,1,K.E.N
A(AIND(K)+K)=A(AIND(K)+K)+1.
SCRAP=A(AIND(K)+L)
A(AIND(K)+L)=A(AIND(K)+N)
DELTA A(AIND(K)+N)=SCRAP
DIAG=A(AIND(1)+1)
THROUGH EPS, FOR J=2,1,J.G.N
EPS A(AIND(1)+J)=A(AIND(1)+J)/DIAG
THROUGH ZETA, FOR J=2,1,J.G.N
THROUGH ETA, FOR I=J,1,I.G.N
SUB=A(AIND(I)+J)
THROUGH IOTA, FOR K=1,1,K.E.J
IOTA SUB=SUB-A(AIND(I)+K)*A(AIND(K)+J)
ETA A(AIND(I)+J)=SUB
DIAG=A(AIND(J)+J)
THROUGH ZETA, FOR I=J+1,1,I.G.N1
SUB=A(AIND(J)+I)
THROUGH LAMBDA, FOR K=1,1,K.E.J
LAMBDA SUB=SUB-A(AIND(J)+K)*A(AIND(K)+I)
ZETA A(AIND(J)+I)=SUB/DIAG
FUNCTION RETURN A(AIND(N)+N1)
END OF FUNCTION
```



## APPENDIX E

### PROGRAM VASYMP TO COMPUTE EQUATION (4.3.13).

RPROGRAM NAME IS VASYMP.

RTHIS PROGRAM RECURSIVELY COMPUTES VALUES OF V(I,J,M) FOR  
RI=1,...,N, J=S,...,T. THE REWARD MATRIX IS R AND THE TERM-  
RINAL REWARD VECTOR IS RHO. A MATRIX BETA PRIOR IS ASSUMED.  
RTHE DISCOUNT FACTOR IS BETA.

```
PROGRAM COMMON R,RHO,RDIM,IND,N,BETA,LIST,TOP
INTEGER N,IND,J,S,T,K,L,TOP
DIMENSION R(100,RDIM),M(100,RDIM),RHO(10),IND(10),V1(10),
OV2(10),LIST(21000)
VECTOR VALUES RDIM=2,1,0
READ READ FORMAT INPUT1, N,S,T,BETA
RDIM(2)=N
IND(1)=0
THROUGH ALFA, FOR K=1,1,K.E.N
ALFA IND(K+1)=K*N
READ READ FORMAT INPUT2, R(1,1) ... R(N,N), M(1,1) ... M(N,N),
ORHO(1) ... RHO(N)
PRINT FORMAT OUT1A, BETA
PRINT FORMAT OUT1B, (K=1,1,K.G.N,(L=1,1,L.G.N, M(IND(K)+L)))
PRINT FORMAT OUT1C,(K=1,1,K.G.N,(L=1,1,L.G.N,R(IND(K)+L)))
PRINT FORMAT OUT1D, (K=1,1,K.G.N, RHO(K))
THROUGH PHI, FOR K=1,1,K.G.N
PHI V2(K)=0.
SET LIST TO LIST
LIST=0
THROUGH DELTA, FOR K=S,1,K.G.T
THROUGH GAMMA, FOR L=1,1,L.G.N
V1(L)=V.(L,K,M)
GAMMA V2(L)=V1(L)-V2(L)
PRINT FORMAT OUT2, K, V1(1) ... V1(N)
PRINT FORMAT OUT3, V2(1) ... V2(N)
THROUGH DELTA, FOR L=1,1, L.G.N
DELTA V2(L)=V1(L)
TRANSFER TO READ

RFORMAT SPECIFICATIONS
VECTOR VALUES INPUT1=$3I10, F10.5*$
VECTOR VALUES INPUT2=$(7F10.5)*$
VECTOR VALUES OUT1A=$8H1BETA  =,G15.5*$
VECTOR VALUES OUT1B=$8H0      M  =,8G15.5/(1H ,8G15.5)*$
VECTOR VALUES OUT1C=$8H0      R  =,8G15.5/(1H ,8G15.5)*$
```



```
VECTOR VALUES OUT1D=$8H0 RHO =,8G15.5/8G15.5*$  
VECTOR VALUES OUT2=$8H0FOR T =, I2, 14H V(I, T, M) =,  
07G15.5/8G15.5*$  
VECTOR VALUES OUT3=$$3,21HDELTA V(I, T=1, M) =,7G15.5/  
08G15.5*$  
END OF PROGRAM
```

EXTERNAL FUNCTION (I1, J1, M)  
ENTRY TO V.

RTHIS FUNCTION RECURSIVELY COMPUTES V.(I1,J1,M)=Y, THE TOTAL  
REXPECTED DISCOUNTED RETURN IN J1 STEPS IF THE SYSTEM STARTS  
RIN STATE I1 WITH PARAMETER MATRIX M. PRIOR IS MATRIX BETA.

```
PROGRAM COMMON R,RHO,RDIM,IND,N,BETA,LIST,TOP  
INTEGER I1,J1,I,J,K,IND,N,RDIM,TOP  
DIMENSION R(100,RDIM),RHO(10),RDIM(2),IND(10),LIST(21000),  
OTM(100,RDIM)  
I=I1  
J=J1  
WHENEVER J .E. 0, FUNCTION RETURN RHO(I)  
MSUM=0.  
THROUGH ALFA, FOR K=1,1,K.G.N  
ALFA MSUM=MSUM+M(IND(I)+K)  
Y=0.  
THROUGH GAMMA, FOR K=1,1,K.G.N  
SAVE RETURN  
SAVE DATA J,Y,MSUM,M(I,K) ... M(I,N),I,K  
EXECUTE TR,(I,K,M,TM)  
X=V.(K,J-1,TM)  
RESTORE DATA K,I,M(I,N) ... M(I,K),MSUM,Y,J  
RESTORE RETURN  
GAMMA Y=Y+(M(IND(I)+K)/MSUM)*(R(IND(I)+K)+BETA*X)  
FUNCTION RETURN Y  
END OF FUNCTION
```



EXTERNAL FUNCTION (I, K, M, TM)  
ENTRY TO TR.

RTHIS FUNCTION EFFECTS THE TRANSFORMATION FROM THE PRIOR  
RPARAMETER MATRIX M TO THE POSTERIOR PARAMETER MATRIX  
 $RT_{\alpha}(I,K,M) = TM$ , WHEN ONE TRANSITION FROM I TO K IS OBSERVED.  
RPRIOR IS MATRIX BETA.

```
PROGRAM COMMON R,RHO,RDIM,IND,N,BETA,LIST,TOP
DIMENSION R(100,RDIM),RHO(10),RDIM(2),IND(10),LIST(21000)
INTEGER I,K,IND,J,L,N
THROUGH ALFA, FOR J=1,1,J,G.N
THROUGH ALFA, FOR L=1,1,L.G.N
ALFA TM(IND(J)+L)=M(IND(J)+L)
TM(IND(I)+K)=TM(IND(I)+K)+1.0
FUNCTION RETURN
END OF FUNCTION
```



## BIBLIOGRAPHY

1. T. W. Anderson and L. A. Goodman, "Statistical inference about Markov chains," *Ann. Math. Stat.* 28 (1957), 89-109.
2. P. Appell and J. Kampé de Fériet. *Fonctions Hypergéométriques et Hypersphériques. Équations Hypersphériques et Hypersphériques.* Gauthier-Villars, Paris (1926).
3. B. Arden, B. Galler, and R. Graham. *The Michigan Algorithm Decoder.* [n.p.] (November, 1963).
4. J. A. Ayres, "Recursive programming in FORTRAN II," *Comm. ACM*, 6 (1963), 667-668.
5. G. A. Barnard, "Sampling inspection and statistical decisions," *J. Roy. Stat. Soc. Ser. B*, 16 (1954), 151-174.
6. R. Bellman. *Adaptive Control Processes: A Guided Tour.* Princeton University Press, Princeton (1961).
7. \_\_\_\_\_, "A problem in the sequential design of experiments," *Sankhyā*, 16 (1956), 221-229.
8. \_\_\_\_\_ and R. Kalaba, "On adaptive control processes," *IRE Trans.* vol. AC-4, no. 2 (1959), pp. 1-9.
9. B. R. Bhat, "Bayes solution of sequential decision problem for Markov dependent observations," *Ann. Math. Stat.* 35 (1964), 1656-1662.
10. P. Billingsley, "Statistical methods in Markov chains," *Ann. Math. Stat.* 32 (1961), 12-40; see also correction in *ibid.*, p. 1343.



11. D. Blackwell, "Discrete dynamic programming," *Ann. Math. Stat.* 33 (1962), 719-726.
12. K. L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer, Berlin (1960).
13. J. M. Cossolino. *Optimal Sequential Decisions Under Uncertainty*. MS Thesis, Massachusetts Institute of Technology (May, 1964).
14. \_\_\_\_\_, R. Gonzalez-Zubieta, and R. L. Miller. *Markovian Decision Processes with Uncertain Transition Probabilities*. Technical Report No. 11, Research in the Control of Complex Systems. Operations Research Center, Massachusetts Institute of Technology (March, 1965).
15. R. Dawson and I. J. Good. "Exact Markov probabilities from oriented linear graphs," *Ann. Math. Stat.* 28 (1957), 946-956.
16. C. Derman, "On sequential decisions and Markov chains," *Biomet. Sci.* 2 (1963), 16-24.
17. A. Erdélyi, et al. *Higher Transcendental Functions*. 3 vols. McGraw-Hill, New York (1953-55).
18. M. Freimer, "A dynamic programming approach to adaptive control processes," *IRE Trans.* vol. AC-4, No. 2 (1959), pp. 10-15.
19. \_\_\_\_\_. *Topics in Dynamic Programming—II. Truncated Policies*. Lincoln Laboratory Report 540-0020. Massachusetts Institute of Technology (18 April 1960).
20. I. J. Good, "The frequency count of a Markov chain and the transition to continuous time," *Ann. Math. Stat.* 32 (1961), 41-48.



31. J. E. Mosimann, "On the compound multinomial distribution, the multivariate  $\beta$ -distribution, and correlations among proportions," *Biometrika*, 49 (1962), 65-82.
32. P. Naur (ed.), "Report on the algorithmic language ALGOL 60," *Comm. ACM*, 3 (1960), 299-314.
33. H. Raiffa and R. Schlaifer. Applied Statistical Decision Theory. Graduate School of Business Administration, Harvard University, Boston (1961).
34. D. Rosenblatt, "On linear models and the graphs of Minkowski-Lorentz matrices," *Econometrica*, 25 (1957), 325-338.
35. W. Rudin. Principles of Mathematical Analysis. 2nd ed. McGraw-Hill, New York (1964).
36. L. S. Shapley, "Stochastic games," *Proc. Nat. Acad. Sci.* 39 (1953), 1095-1100.
37. N. Z. Shor, "Pro optimal'ne reguliuvaniya Markova'koj poslidovnosti z dvema fazovimi stanazi" (On the optimal control of a Markov chain with two phase states; Ukrainian, with Russian summary), *Zbirnik Prats' z Obchisllyval'imi Matematicheskimi i Tekhnicheskimi Akademiym Nauk URSR*, Kifiv, 1 (1961), 119-124.
38. E. A. Silver. Markovian Decision Processes with Uncertain Transition Probabilities or Rewards. Technical Report No. 1, Research in the Control of Complex Systems. Operations Research Center, Massachusetts Institute of Technology (August, 1963).



39. A. Singer, "The steady state probabilities of a Markov chain as a function of the transition probabilities," *Qua. Reg.* 12 (1964), 498-499.
40. G. B. Wetherill, "Bayesian sequential analysis," *Biometrika*, 48 (1961), 281-292.
41. P. Whittle, "Some distribution and moment formulae for the Markov chain," *J. Roy. Stat. Soc. Ser. B* 17 (1955), 235-242.
42. L. E. Zachrisson, "Markov games," in *Advances in Game Theory*, M. Dresher, L. S. Shapley, and A. W. Tucker, eds. Princeton University Press, Princeton (1964). pp. 211-253.



## BIOGRAPHICAL NOTE

James J. Martin, Jr. was born on February 3, 1936 in Paterson, New Jersey, where he attended elementary school and Eastside High School. In 1951, after completing three years of high school, he entered the University of Wisconsin, where he was supported by a Ford Foundation scholarship. In June, 1955, he was awarded a Bachelor of Arts degree in Physics by the University of Wisconsin and won the Vilas Prize Essay Contest.

Mr. Martin studied Theology at the Harvard Divinity School from 1955 to 1957. In July, 1957, he graduated from Officer Candidate School and was commissioned an Ensign in the U. S. Navy Reserve. Mr. Martin is a career Naval officer on active duty and currently holds the rank of Lieutenant. He has served as Executive Officer in USS BATTLEBORN (EPCER 852) and as Engineer Officer in USS VESOLE (DDR 878). From 1961 to 1963, Mr. Martin attended the U. S. Naval Postgraduate School, where he studied Operations Research, receiving the degree of Master of Science in May, 1963. He has been enrolled in the graduate school of the Massachusetts Institute of Technology since February, 1963.

Mr. Martin's publications include "Multinormal Bayesian Analysis: Two Examples," published by the Sloan School of Management of M.I.T., and "Distribution of the Time Through a Directed, Acyclic Network," published in Operations Research. He presented a paper entitled "On the Expected Gain of a Markov Chain with Uncertain Transition Probabilities" at the Twenty-Seventh National Meeting of the Operations Research Society of America.

Mr. Martin is a member of the Tau Beta Pi and Sigma Xi honor societies and belongs to the Operations Research Society of America.

In 1954, Mr. Martin was married to the former Miss Betty Bent of Benton, Wisconsin. They now have four children.





thesM3576

Some Bayesian decision problems in a Mar



3 2768 002 12817 5

DUDLEY KNOX LIBRARY